

GenCore version 4.5
 Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on:

May 19, 2002, 08:06:58 ; Search time 67.41 Seconds

(without alignments)

1255.574 Million cell updates/sec

Title:

US-09-673-302a-1

Perfect score:

4154

Sequence:

1 GNCITTRGVSSCQQCLAVS.....NNPLYKEATSTFTNTIYGT 762

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched:

747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters:

747574

Minimum DB seq length:

0

Maximum DB seq length:

200000000

Post-processing:

Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Genesed_032802:*

1: /SIDS5/gcdata/geneseq/geneseqp-emb1/AA1980.DAT:*

2: /SIDS5/gcdata/geneseq/geneseqp-emb1/AA1981.DAT:*

3: /SIDS5/gcdata/geneseq/geneseqp-emb1/AA1982.DAT:*

4: /SIDS5/gcdata/geneseq/geneseqp-emb1/AA1983.DAT:*

5: /SIDS5/gcdata/geneseq/geneseqp-emb1/AA1984.DAT:*

6: /SIDS5/gcdata/geneseq/geneseqp-emb1/AA1985.DAT:*

7: /SIDS5/gcdata/geneseq/geneseqp-emb1/AA1986.DAT:*

8: /SIDS5/gcdata/geneseq/geneseqp-emb1/AA1987.DAT:*

9: /SIDS5/gcdata/geneseq/geneseqp-emb1/AA1988.DAT:*

10: /SIDS5/gcdata/geneseq/geneseqp-emb1/AA1989.DAT:*

11: /SIDS5/gcdata/geneseq/geneseqp-emb1/AA1990.DAT:*

12: /SIDS5/gcdata/geneseq/geneseqp-emb1/AA1991.DAT:*

13: /SIDS5/gcdata/geneseq/geneseqp-emb1/AA1992.DAT:*

14: /SIDS5/gcdata/geneseq/geneseqp-emb1/AA1993.DAT:*

15: /SIDS5/gcdata/geneseq/geneseqp-emb1/AA1994.DAT:*

16: /SIDS5/gcdata/geneseq/geneseqp-emb1/AA1995.DAT:*

17: /SIDS5/gcdata/geneseq/geneseqp-emb1/AA1996.DAT:*

18: /SIDS5/gcdata/geneseq/geneseqp-emb1/AA1997.DAT:*

19: /SIDS5/gcdata/geneseq/geneseqp-emb1/AA1998.DAT:*

20: /SIDS5/gcdata/geneseq/geneseqp-emb1/AA2000.DAT:*

21: /SIDS5/gcdata/geneseq/geneseqp-emb1/AA2001.DAT:*

22: /SIDS5/gcdata/geneseq/geneseqp-emb1/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	4119	99.2	788	20 AAY4953 Human endothelial
2	4119	99.2	788	20 AAY49567 Human antithrombin
3	4119	99.2	788	21 AAY92413 Wild type human GP
4	4119	99.2	788	21 AAY8161 Human integrin bet
5	4115	99.1	762	21 AAY92451 GPIIa variant Arg
6	4115	99.1	788	15 AAY5150 Human platelet GPI
7	4114	99.0	762	21 AAY92456 GPIIIa variant Ser
8	4113	99.0	762	21 AAY92448 GPIIIa variant Leu
9	4113	99.0	762	21 AAY92444 GPIIIa variant Ser
10	4112	99.0	788	21 AAY92444 Variant human GPII
11	4111	99.0	762	21 AAY92452 Variant human Arg

ALIGNMENTS

RESULT 1	AY4953	AY4953 standard; Protein; 788 AA.
	XX	XX
	AC	AY4953;
13-JAN-2000	XX	(first entry)
	XX	
	DE	Human endothelial membrane glycoprotein IIIa protein sequence.
	KW	Human; coding sequence polymorphism; vascular pathology gene; polymorphic site; phenotype correlation; forensic; paternity testing; medicine; genetic analysis; vascular disease.
	KW	
	OS	Homo sapiens.
	XX	
	PN	W09950454-A2.
	XX	
	PD	07-OCT-1999.
	XX	
	PR	26-MAR-1999; 99W0-US06473.
	XX	
	PR	01-APR-1998; 98US-0054272.
	XX	
	PA	(WHED) WHITEHEAD INST BIOMEDICAL RES.
	XX	
	PI	Lander ES, Daley GQ, Cargill M, Ireland JS, Rozen SG;
	XX	
	DR	WPI; 1999-620066/53.
	DR	N-PSB; AAZ32162.
	XX	
	PT	Determination of polymorphisms in genes, especially those identifying predisposition to vascular disease
	XX	

GPIIIa variant Pro
 GPIIIa variant Arg
 GPIIIa variant Asp
 GPIIIa variant Cys
 Mouse beta-3 integrin
 Secreted GPIIIa subunit
 Novel human extracellular matrix protein
 Mouse integrin beta-3
 Human integrin bet
 Porcine CPG prote
 Human polypeptide
 Guinea pig integrin
 Beta-1 integrin subunit
 Human CD18 for use
 Rabbit beta-1 integrin
 Human beta-8 integrin
 Sequence of a peptide
 Sequence of a peptide

PS Disclosure: Fig 6; 134pp; English.

XX AA232159 to AA232194 represent reference alleles for specifically claimed nucleic acid sequences from the present invention which comprise polymorphic sites as given in a table in the specification, selected from 92 single nucleotide polymorphisms in which the nucleotide at the polymorphic site is different from a nucleotide at the same site in a reference allele. The nucleic acids, and primers and probes, are used to identify polymorphisms, which may predispose an individual to disease, especially a vascular disease. They can also be used in phenotype correlations, forensics, paternity testing, medicine or genetic analysis.

CC AAV49550 to AAV49573 represent the proteins which correspond to some of the reference alleles.

XX Sequence 788 AA;

SQ

Query Match

Best Local Similarity

99.2%; Score 4119; DB 20; Length 788;

Matches 757; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 GPNICTRGSSCQCLAVSPMCACWCSDEALPLGSPRCDLKENLKCAPESIEFPVSE 60

Db 27 gpnicttggvssccqclavspmcacwscdealplgsprcdlkenlkcapesiefpvse 86

QY 61 ARVLEDRPLDKSGDSSQTVQSPORALRRLRPPDSKNSFIQVRQVEDPYVIVYMLD 120

Db 87 arviledpldksgdssqtvqspqrallrpddsknfisqvrqvedpyvivyml 146

QY 121 SYSMKDDLWIONLQGTLKATOMRKLTSNLRIGFAFVVKPVSPYMYISPPPEALENPVCDM 180

Db 147 sysmkddlwsinqlgktlkatqmkltsnlrlqfqaafvdkppspymisppalepnydm 206

QY 181 KTTCLPMPGKVKHVLTLTDQTVRNEEVKKQSVSNRDAPEGGDAMQATVDEKIGWRN 240

Db 207 kttclpmgkvhvltldqtrnneevkkqsvsnrdapeggdaimatvdekeigw 266

QY 241 DASHLVLVFTDAKTHIALDGLLAGIVOPNDGOCHVGSNDHSASTMDYPSIGLMTKLS 300

Db 267 dashliveltdakthialdgrlagivqpgndqgchvgsdnhyasttdyplsgimtekis 326

QY 301 OKNINLIAVTEVNVNLYQNSLIPGTVGVLMSMDNSNVNOLIVDAYKGIRSKVLEVR 360

Db 327 qkninlifavtenevnlyqnslyeipgtgvylsmdssnvnqlivdaykgirkvlevr 386

QY 361 DLPEELSLSPNATCLNNENVIPGKSCMGLKIGDTVSTIESKAVRGCPQEKEKSTIKPGV 420

Db 387 dlpeelslsfnaclnnevpgikscmglikgtvtsiseakvrgcpgekstikpgv 446

QY 421 FKPSLIVQVTFDCDCACQAAQEPNSHRCNNNGNTFECGVRCGGPGWLGSOCSESDYRP 480

Db 447 fkaslivavtfdcacqaaqepnshrcnnngntfecgvrcggpgwlgsocseedyrp 506

QY 481 SQDDECSREGQVCSQRGECLGQCVCHSSPFGKTKYECDDSCVRYKGEMSGHG 540

Db 507 sqddecspregqvcsgqrgeclgqcvchssafgkftgkyceddfscvrykgemsghg 566

QY 541 QCSCGDCLCSDMWGTYCNCNTTRDTONMSNLCSRGKFCGSCVCIOQPSYGTCEK 600

Db 567 qscgdcldcsdwtgycncttrdtonmsnlcsrgkfcgscvcipgsgyatek 626

QY 601 CPTCPDACTFKKCVCHECKFDRGALHDENTCRYCREDEIESVKELKDTGKAVNCTYKE 660

Db 627 cptcpdaactfkkeveckfdrepymtentchrycdeiesvkelkdtgkavnctyke 686

QY 661 DDCWVRYQYESSGSKLIVVVEPECPKGPDILWILSVAGAILIGLALLKLIT 720

Db 687 ddcwvryqyessgsklilyvvepecpkgpdilwilsngailiglaalliklit 746

QY 721 IHRKEKEFEERRAKWDTANPLYKEAESTFTNTYRST 762

Db 747 ihrkekefeerrakwdtanplykeastftntyrgt 788

RESULT 2

AAV49567

ID AAV49567 standard; Protein: 788 AA.

XX

AC AAV49567;

XX

DT 13-JAN-2000 (first entry)

XX DE Human antithrombin III protein sequence.

XX KW Human; coding sequence polymorphism; vascular pathology gene;

KW polymorphic site; phenotype correlation; forensic; paternity testing; medicine; genetic analysis; vascular disease.

XX

OS Homo sapiens.

XX

PN W0950454-A2.

XX

PD 07-OCT-1999.

XX

PF 26-MAR-1999; 99WO-US06473.

XX

PR 01-APR-1998; 98US-0054272.

XX

PA (WHED) WHITEHEAD INST BIOMEDICAL RES.

XX

PI Lander ES, Dailey GO, Cargill M, Ireland JS, Rozen SG;

XX

DR WEI; 1999-620066/53.

N-PSDB: AA232186.

XX

DR AA232186.

XX

PT Determination of polymorphisms in genes, especially those identifying

PT predisposition to vascular disease

PS Disclosure: Fig 30; 134pp; English.

XX

AA232159 to AA232194 represent reference alleles for specifically claimed nucleic acid sequences from the present invention which comprise polymorphic sites as given in a table in the specification, selected from 92 single nucleotide polymorphisms in which the nucleotide at the polymorphic site is different from a nucleotide at the same site in a reference allele. The nucleic acids, and primers and probes, are used to identify polymorphisms, which may predispose an individual to disease, especially a vascular disease. They can also be used in phenotype

correlations, forensics, paternity testing, medicine or genetic analysis.

AAV49550 to AAV49573 represent the proteins which correspond to some of the reference alleles.

XX

SQ Sequence 788 AA;

Query Match

Best Local Similarity

99.2%; Score 4119; DB 20; Length 788;

Matches 757; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 GPNICTRGSSCQCLAVSPMCACWCSDEALPLGSPRCDLKENLKCAPESIEFPVSE 60

Db 27 gpnicttggvssccqclavspmcacwscdealplgsprcdlkenlkcapesiefpvse 86

QY 61 ARVLEDRPLDKSGDSSQTVQSPORALRRLRPPDSKNSFIQVRQVEDPYVIVYMLD 120

Db 87 arviledpldksgdssqtvqspqrallrpddsknfisqvrqvedpyvivyml 146

QY 121 SYSMKDDLWIONLQGTLKATOMRKLTSNLRIGFAFVVKPVSPYMYISPPPEALENPVCDM 180

Db 147 sysmkddlwsinqlgktlkatqmkltsnlrlqfqaafvdkppspymisppalepnydm 206

QY 181 KTTCLPMPGKVKHVLTLTDQTVRNEEVKKQSVSNRDAPEGGDAMQATVDEKIGWRN 240

Db 207 kttclpmgkvhvltldqtrnneevkkqsvsnrdapeggdaimatvdekeigw 266

QY 241 DASHLVLVFTDAKTHIALDGLLAGIVOPNDGOCHVGSNDHSASTMDYPSIGLMTKLS 300

XX OS Homo sapiens.
 XX OS WO200020634-A1.
 XX PN 13-APR-2000.
 XX PD 01-OCT-1999; 99WO-IB01696.
 XX PR 01-OCT-1998; 98US-0102624.
 XX PA (NOVA-) NOVA MOLECULAR INC.
 XX PT Schappert, K.;
 XX DR WPI; 2000-303801/26.
 XX PT Identifying a subject at risk for a neurological disease comprises determining of genotype or phenotype of GPIIa or GPIIb locus and determining presence of variant GPIIa or GPIIb allele or isoform
 XX PS Disclosure; Page -; 55pp; English.
 CC Known polymorphisms in GPIIa that may be determined to be variants using the methods of the invention include AAY2447-56, which are variant GPIIa sequences derived from the mature protein.
 CC Identifying a subject at risk for a neurological disease comprises determining the genotype or phenotype of the GPIIa or GPIIb locus of a subject, and determining the presence of a variant GPIIa or GPIIb allele or isoform, where the presence indicates an increased risk of neurological disease. The method is useful for identifying a risk, diagnosing and treating a neurological disease comprising Alzheimer's disease, neurofibromatosis, Huntington's disease, depression, amyotrophic lateral sclerosis, multiple sclerosis, stroke, Parkinson's disease and multi-infarct dementia.
 CC Note: This sequence is not given in the specification it was created from the wild type human GPIIa sequence which appears in Figure 3.
 XX SQ Sequence 762 AA:

Query Match 99.1%; Score 4115; DB 21; Length 762;
 Best Local Similarity 99.2%; pred. No. 1.26-282; 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 GPNICTTRGVSSCQCLAVSPMCACWSDEALPAGSPRODKEENLKLNCAPSEIEFPYE 60
 Db 1 gpnlcttrgvsscqcavspmcacwsdealpagsprodkeenlkncapseiaypse 60

QY 61 ARVHEDRPLSDKGSDSSQVTOYSPORITALRRLRPPDKNSKFSTORQVEDPYDYYLMDL 120
 Db 61 arviedrplsdkgsgdssqvtqyvpsqprairlrrpdksnfslqrqvdyypdyiyldl 120

QY 121 SYSKKDLQSIONGKTLATOMRKLTURIGRIGRAFVDPKPVSYWYMSPPALENPYDM 180
 Db 121 sysnkdlwsiqnlqtklqmktsnlrlqgfvdkpvspymysppalelpvdm 180

QY 181 KTCICLPMFGKVHVLTDQVTRNEEVKQSVSRNDAPEGGDAIMQATVDEKEIGHN 240
 Db 181 kttciclpmpfkyhvtldqytfiefveekqqsqsnrdapeggdaimqatvdekgwn 240

QY 241 DASHLVLFTDAKTHIALDGRLAGIVOPNGOCHVGNSDHYASSTTMDYPSLGLMTKLS 300
 Db 241 dashlvlftdakthialdgrlagivopngqchvgnsdhyasttmdyplsgimteklis 300

QY 301 QKNTNLIFAVTENYVNLQYNSSEIPIGTVGVLSSMDSSWVQIIVDAYGKISKVELEVR 360
 Db 301 qknntnlifavtenvnlqyngseiplgtpgvmlsmdsswvqivdaygkiskvelevr 360

QY 361 DLPEELSLSENATCINNEYTPIGLSCMGKIKGTVMSFSTEAKVKGCPOEKESFTIKPGV 420
 Db 361 dlpeelslsfnatcinnneytppgkscmgkikgtsfsleakrgcpkkesftikpgv 420

XX FH RESULT 6
 FT ID AAR51500 standard; Protein: 788 AA.
 FT XX AAR51500;
 FT XX DT 17-NOV-1994 (first entry)
 FT XX DE Human platelet GPIIa surface antigen Yuk-a.
 FT XX KW platelet surface antigen; glycoprotein GPIIa; Yuk antigen; probe;
 FT XX OS detection; thrombocyte; Yuk-a allele.
 XX FH Key Location/Qualifiers
 FT Misc-difference 169 /note= "Yuk-a and Yuk-b differ at position 169 w/with
 FT Gln (corresp. to CAA codon) in Yuk-a but
 FT Arg (corresp. to CGA codon) in Yuk-b"
 FT XX PR 06-JUL-1991; 91JP-0192554.
 FT PR 06-JUL-1991; 91JP-0192554.
 FT XX PA (JUJI/) JUJI T.
 FT XX DR WPI; 1994-131276/16.
 FT DR N-PSDB; AA062199.
 XX PT The detection of platelet surface antigen determinant - using
 XX PT oligo-nucleotide probes for Yuk a and Yuk b
 XX Disclosure; Fig 1-5 (Page 8-12); 12pp; Japanese.
 CC Primers (3) and (4) (AA062195 and AA062196, respectively) were used to amplify a region of the Yuk gene from platelet cDNA derived from humans of known Yuk antigen type. The 1mer probes AA02193 and AA062194, specific for the Yuk a and Yuk b alleles respectively, were tested for hybridisation to the cDNA. The Yuk-a probe hybridised to cDNA from a Yuk a/b homozygote and from Yuk a/b heterozygotes only, while the Yuk b probe hybridised to cDNA from the heterozygotes and from Yuk b/b homozygotes only. AAR51500 is the sequence for the human Yuk-a protein.
 CC

XX	XX	KW	neurofibromatosis; Huntingdons disease; depression; Parkinsons disease;
SQ	Sequence	KW	anyotrophic lateral sclerosis; multiple sclerosis; dementia.
Query Match	99.1%; Score 4115; DB 15; Length 788;		
Best Local Similarity	99.2%; Pred. No. 1; e-282;		
Matches	756; Conservative 1; Mismatches 5; Indels 0; Gaps 0;		
QY	1 GPNICTTREVSSCOOCLAVSPMCWCSDDEALPLIGSPRDKLKENLKDNCAPBESIEFPVSE 60	OS	Homo sapiens.
Db	27 gpniicttgrvsscqclavspmcawcsdealplgsprcdklenlkdcapiesiefpse 86	XX	
QY	61 ARVLEDRLSDKSGDSSQTOVSPORALRPLRDPDKNSFQVROQVEDYDVIYLMQL 120	OS	
Db	87 arvledrlsdksqgssqvtpqspqralrplrdksnfqvrqvedydpdyimdl 146	XX	
QY	121 SYSMKDDLWSIQNLGKLTQMRKLTSLNRIGFAVFKPSPIMIISPPALENPYDM 180	OS	
Db	147 syamkddlwsiqplgklatqmktsnlrlqfagafvdksprcdklenlkdcapiesiefpse 206	XX	
QY	181 KTCCLPMGKHYHLTDQVTRENEEVKQSVSRNRDAPEGFDATMQATVCDEKIGWRN 240	PA	(NOVA-) NOVA MOLECULAR INC.
Db	207 ktclpmfgkvnltldqvtifnevkqsvsrndapeqgfdaimqatvcdkewrn 266	XX	
QY	241 DASHLIVFTTDKTHIALDGLAGIVOPNDSDQCHGSDNHYASTMDYSSGLMTEKL 300	DR	WPI: 2000-303801/26.
Db	267 dashliffttdkthialdgrlagivqndqgchvsdnhyastmdyssglmtekl 326	XX	
QY	301 OKNINLIFAVTENVNQYQNTSELICGTVVLSMOSNVQQLIVDAYGIRSKVELER 360	PT	Identifying a subject at risk for a neurological disease comprises
Db	327 qkniniifavtenvnlyqnselipggtvylsmdssnvqilividaygirkvelevr 386	PT	determination of genotype or phenotype of GPIIa or GPIb locus and
QY	361 DPEELSLSFNATCLNNEVICPLKSCMGKIGDTSFSIIRAKVRGCPQEKEKSFITKPG 420	PT	determining presence of variant GPIIa or GPIb allele or isoform
Db	387 dpeelslshatclnnenvicplkscmgkigdtsfsieakvrgcpqeksfktpg 446	XX	
QY	421 FRSILIVQVTFDCACQAEKAPBNHRCNNNGTFCGVERGCPGWLGSCECSDEYDYP 480	PS	Disclosure: Page -; 55PP; English.
Db	447 fksilivqvtfdcdacqacqapepnshrcnnngtfcgvcrgcpwlgsgcseedyyp 506	CC	Known polymorphisms in GPIIa that may be determined to be variants
QY	481 SQQEDESPREPOVPCSORGECLGCGVCHSSDFGKLTGKVCCECDESCVYKGERCGH 540	CC	using the methods of the invention include AAY2447-56, which are variant
Db	507 sqgdecspreqpvcsqrgcgeclcgqvcgchssdfgktykrecadfcvrvkqkgeomsgng 566	CC	GPIIa sequences derived from the mature protein.
QY	541 QSCCGDLCDCSDWTGTYCNCNTTRDTMCMSSNGLCOSGRGKCECGSCVCLIQPGSVDTCEK 600	CC	Identifying a subject at risk for a neurological disease comprises
Db	567 qscscgdcldsdwtgycncnttrdtcmssnqllsosgrgkcecgscvcqcgpsqdtcek 626	CC	determining the genotype or phenotype of the GPIIa or GPIb locus of a
QY	601 CPTCPCACTFKRKEVCKKFERGALDENTCNCYRDELESVKELDTGDAVNTYKE 660	CC	subject, and determining the presence of a variant GPIIa or GPIb
Db	627 optcpdactfkkevckkfergatntcncrydelesvkelkdtgdkavntyke 686	CC	allele or isoform, where the presence indicates an increased risk of
QY	661 DDCVWRQYDSSGSKILVYVEEECPKRPDILVLLSTMAGATLIGAALLIWKLIT 720	CC	neurological disease. The method is useful for identifying a risk,
Db	687 ddcvwrqyddssgskilvyyveebeepckpkdpdilvllsmaigalliaalliwklit 746	CC	diagnosing and treating a neurological disease comprising Alzheimer's
QY	721 KDRKIFAKFEKFEERAKWMTANNLYKEISTSTFNITRGT 762	CC	disease, neurofibromatosis, Huntingdon's disease, depression,
Db	747 ldrkefakfeerakwmtannlykeatstfnityrgt 788	CC	disease and multi-infarct dementia.
QY	7 Sequence 762 AA;	XX	Note: This sequence is not given in the specification it was created
QY	7 Sequence 762 AA;	XX	from the wild type human GPIIa sequence, which appears in Figure 3.
Query Match	99.0%; Score 4114; DB 21; Length 762;		
Best Local Similarity	99.2%; Pred. No. 1; 4e-282;		
Matches	756; Conservative 0; Mismatches 6; Indels 0; Gaps 0;		
QY	1 GPNICTTREVSSCOOCLAVSPMCWCSDDEALPLIGSPRDKLKENLKDNCAPBESIEFPVSE 60		
Db	1 gpniicttgrvsscqclavspmcawcsdealplgsprcdklenlkdcapiesiefpse 86		
QY	61 ARVLEDRLSDKSGDSSQTOVSPORALRPLRDPDKNSFQVROQVEDYDVIYLMQL 120		
Db	61 arvledrlsdksqgssqvtpqspqralrplrdksnfqvrqvedydpdyimdl 146		
QY	121 SYSMKDDLWSIQNLGKLTQMRKLTSLNRIGFAVFKPSPIMIISPPALENPYDM 180		
Db	121 syamkddlwsiqplgklatqmktsnlrlqfagafvdksprcdklenlkdcapiesiefpse 206		
QY	181 KTCCLPMGKHYHLTDQVTRENEEVKQSVSRNRDAPEGFDATMQATVCDEKIGWRN 240		
Db	181 ktclpmfgkvnltldqvtifnevkqsvsrndapeqgfdaimqatvcdkewrn 266		
QY	241 DASHLIVFTTDKTHIALDGLAGIVOPNDSDQCHGSDNHYASTMDYSSGLMTEKL 300		
Db	301 qkniniifavtenvnlyqnselipggtvylsmdssnvqilividaygirkvelevr 360		
QY	361 DPEELSLSFNATCLNNEVICPLKSCMGKIGDTSFSIIRAKVRGCPQEKEKSFITKPG 420		

Db 361 dpeelslfnatlnneviplgikscmglikatvsfieakrgcpaekstifpvq 420
 CC disease, neurofibromatosis, Huntington's disease, depression,
 CC amyotrophic lateral sclerosis, multiple sclerosis, stroke, Parkinson's
 CC disease and multi-infarct dementia.
 Note: This sequence is not given in the specification it was created
 from the wild type human GpIIa sequence which appears in Figure 3.

Db 421 FKSILIVQTFCDCACAOAQEPNSHRENNGNTFECVVCRCRGWLGSOQCSSEEDYRP 480
 CC
 Db 421 fksilivqtfcdcacacqagaaepnshrcnngntfecqvcrcrgwlgsgqecseedyrp 480
 CC
 Qy 481 SQODECSPREGOPVCSORGECLCQCVCHSSDPKGKIGKYCEGSCDFSCVRYKEMCSHG 540
 CC
 Db 481 sqgdecspregopvcsorgeclcqcvchssdpgkgkigkycegscdfscvrykemcsghg 540
 CC
 Qy 541 QCSGDCLCDSWVGYCVCNTTRDTMSSNGLCSGRKCEGSCVCIOPGSIGEMCSHG 600
 CC
 Db 541 qcsgdclcdswhgqygnctrtdtmsnglicsgrgkigkycegscvqpgsydtcek 600
 CC
 Qy 601 CPTCPDACTFKKECKEVEKKFDRGALHDENTCNRCRDETESVKELKDGKDAVNCTYNE 660
 CC
 Db 601 cptcpdactfkkeckfkdrpmtyntcnrcdelesvkeldtgkdvavntcyne 660
 CC
 Qy 661 DDCVVRFOYEDSGSKSLIVVVEPECPKGPDILWVLSVMAGLLGLAIIWKLIT 720
 CC
 Db 661 ddvvrfoyyedsgkslilyvvepecpkgpdilwvlsvmaigliaglialliwklit 720
 CC
 Qy 721 IHDKEFAKEERARAKWDTANPPLYKEATSTFTNTYRGT 762
 CC
 Db 721 ihorkefakfeerarakwdtannpplyleatstftnityrgt 762
 CC

RESULT 8

AY92448 standard; Protein: 762 AA.

ID AY92448

AC

XX

XX 10-AUG-2000 (first entry)

XX DE GPIIa variant Leu17Trp.

XX

XX GPIIa; GPIb; neurological disease; genotype; phenotype; diagnosis; alantigen; platelet antigen; PAl; PI2; Alzheimer's disease; stroke; neurofibromatosis; Huntington's disease; depression; Parkinson's disease; amyotrophic lateral sclerosis; multiple sclerosis; dementia; Homo sapiens.

XX WO200020634-A1.

XX 13-APR-2000.

XX 01-OCT-1999; 99WO-1B01696.

XX 01-OCT-1998; 98US-0104624.

PA (NOVA-) NOVA MOLECULAR INC.

XX

PT Schappert K;

DR WPI; 2000-303801/26.

XX

PT Identifying a subject at risk for a neurological disease comprises determining of genotype or phenotype of GpIIa or GPIb locus and determining presence of variant GpIIa or GPIb allele or isoform Disclosure; Page - ; 55pp; English.

CC Known polymorphisms in GpIIa that may be determined to be variants using the methods of the invention include AY92447-56, which are variant GpIIa sequences derived from the mature protein.

CC Identifying a subject at risk for a neurological disease comprises determining the genotype or phenotype of the GpIIa or GPIb locus of a subject, and determining the presence of a variant GpIIa or GPIb allele or isoform, where the presence indicates an increased risk of neurological disease. The method is useful for identifying a risk, diagnosing and treating a neurological disease comprising Alzheimer's

CC disease, neurofibromatosis, Huntington's disease, depression, amyotrophic lateral sclerosis, multiple sclerosis, stroke, Parkinson's disease and multi-infarct dementia.

CC Note: This sequence is not given in the specification it was created from the wild type human GpIIa sequence which appears in Figure 3.

XX Sequence 762 AA;

Query Match 99.0%; Score 4113; DB 21; Length 762;

Best Local Similarity 99.2%; Pred. No. 1. 6e-282; Mismatches 6; Indels 0; Gaps 0;

Matches 756; Conservative 1

Db 1 gpiictqtrgqsvscqgqclavspmcawsdealplgsprcdkkenlkdnccpselefpsv 60

Qy 61 ARYLEDRPLSKGSQGSSQVNTQVSPORIALTRPDQSKNTSIQVQROVEDYPVDIYMLD 120

Db 61 aryledrplskgsqgssqvqspqrailrpdknsisqvqvedyvdiyml 120

Qy 121 SISMKDDIWLSTONLGRKLTATRKKLNSLRGFGAIVDVKPSPYMSIPRBALEPCYD 180

Db 121 sismkddiwlstnlgkltatrkknlsrgrgafdkpkspsymisprbaalepcydm 180

Qy 181 KTCCLPARGYKHLTEPDQVRFNEEVKKQSVSNRDAPECGRDAMQATWCDEKIGWRN 240

Db 181 ktcclpmfgqykhvltidqtrtneevkqgqsvsnrdqgqfdaimgatwcdkekgwn 240

Qy 241 DASHLILVFTTDAKTHALDGRLAGITQVQPNQDNGOCHVQSDNHYSASITMDYSSGLMTEKLS 300

Db 241 dashlilvfttakthaldgrlagitqvqpnqdnqgqhsnhsastmdyssglmtekls 300

Qy 301 QKINLTFAVTENVVNYQNTSELIPPTVGLSMQSSNVNQLVDAYGKTRSKVBLEVR 360

Db 301 qknliflafavtenvnyqnselqnttvaylsmassnnyqilividaygrtskelevr 360

Qy 361 DPEELSLSFNTCLNNEVINGLKSMLGKIDTVPSIEAKVKGCPUEKEKSTIKPVG 420

Db 361 dpeelslfnatlnneviplgikscmglikatvsfieakrgcpaekstifpvq 420

Qy 421 FKSILIVQTFCDCACAOAQEPNSHRENNGNTFECVVCRCRGWLGSOQCSSEEDYRP 480

Db 421 fksilivqtfcdcacacqagaaepnshrcnngntfecqvcrcrgwlgsgqecseedyrp 480

Qy 481 SQODECSPREGOPVCSORGECLCQCVCHSSDPKGKIGKYCEGSCDFSCVRYKEMCSHG 540

Db 481 sqgdecspregopvcsorgeclcqcvchssdpgkgkigkycegscdfscvrykemcsghg 540

Qy 541 QCSGDCLCDSWVGYCVCNTTRDTMSSNGLCSGRKCEGSCVCIOPGSIGEMCSHG 600

Db 541 qcsgdclcdswhgqygnctrtdtmsnglicsgrgkigkycegscvqpgsydtcek 600

Qy 601 CPTCPDACTFKKECKEVEKKFDRGALHDENTCNRCRDEIEVEKELDGKAVNTYNE 660

Db 601 cptcpdactfkkeckfkdrpmtyntcnrcdelesvkeldtgkdvavntcyne 660

Qy 661 DDCVVRFOYEDSGSKSLIVVVEPECPKGPDILWVLSVMAGLLGLAIIWKLIT 720

Db 661 ddvvrfoyyedsgkslilyvvepecpkgpdilwvlsvmaigliaglialliwklit 720

Qy 721 IHDKEFAKEERARAKWDTANPPLYKEATSTFTNTYRGT 762

Db 721 ihorkefakfeerarakwdtannpplyleatstftnityrgt 762

CC

RESULT 9

AY92450

ID AY92450 standard; Protein: 762 AA.

XX

AC AY92450;

XX

DT 10-AUG-2000 (first entry)

XX
DE
GPIIIa variant Ser162Leu.
XX
KW
GPIIIa; GPIIB; neurological disease; genotype; phenotype; diagnosis;
alloantigen; platelet antigen; PIA1; PLA2; Alzheimer's disease; stroke;
KW
neurofibromatosis; Huntington's disease; depression; Parkinson's disease;
XX
amyotrophic lateral sclerosis; multiple sclerosis; dementia.
OS
Homo sapiens.
XX
PN
WO200020634-A1.
XX
PD
13-APR-2000.
XX
PF
01-OCT-1999; 99WO-IB01696.
XX
PR
01-OCT-1998; 98US-0102624.
XX
PA
(NOVA-) NOVA MOLECULAR INC.
XX
PI
Schappert K;
XX
DR
WPI; 2000-303801/26.
XX
PT
Identifying a subject at risk for a neurological disease comprises
determination of genotype or phenotype of GPIIIa or GPIIB allele or isoform
determining presence of variant GPIIIa or GPIIB allele or isoform
XX
PS
Known polymorphisms in GPIIIa that may be determined to be variants
using the methods of the invention include AAY92447-56, which are variant
GPIIIa sequences derived from the mature protein.
CC
Identifying a subject at risk for a neurological disease comprises
determining the genotype or phenotype of the GPIIIa or GPIIB locus of a
subject, and determining the presence of a variant GPIIIa or GPIIB
allele or isoform, where the presence indicates an increased risk of
neurological disease. The method is useful for identifying a risk,
diagnosing and treating a neurological disease comprising Alzheimer's
disease, neurofibromatosis, Huntington's disease, depression,
CC
amyotrophic lateral sclerosis, multiple sclerosis, stroke, Parkinson's
disease and multi-infarct dementia.
Note: This sequence is not given in the specification it was created
from the wild type human GPIIIa sequence which appears in Figure 3.
XX
Sequence 762 AA:

Query Match 99.0%; Score 4113; DB 21; Length 762;
Best Local Similarity 99.2%; Pred. No. 1 6e-202; Gaps 0;
Matches 756; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GPNICTRTGVSQSCQCLAVSPKMCAWSDEALPLGSPRCPLKLENLKDNCAPESIEFPVSE 60
Db 1 gpniicttggvssccqclavspkmcawsdealplgsbrcdlklenlkdcapiesiefpse 60
QY 61 ARVLEDPLDKSGDSSQVTOYSPQRLALRURPDDSKNFESIQVROVEDYPVDIYMLD 120
Db 61 arvledpldksgdssqvtqspqrlalrurpddsknfesiqvrvqvedypvdiyml 120
QY 121 SISMKDDIWSNGLGKIAKTRKRLKTSNLRGAFVDPKSPYMTSPBEPALENPCDM 180
Db 121 sysmkddlwsgnlgktklatqmkrlsnrlrgfgrfvdkpypymispealencyd 180
QY 181 KTCCLPLPHGPKVLTIDQVRFNEEVKVKQSVRSNDAPECGFDAQMAWCDEKIGWRN 240
Db 181 ktcclplphgpkvltidqvrfrneevkqksvrsnrdapecgfdalmqatvrdkewr 240
QY 241 DASHLYVFTTAKTHALDGRLAGIQPNQDGSQCHVGSNDHYSASTMDYFSSGLMTEKLS 300
Db 241 dashlyvfttakthaldgrlagiqvpndqgchqsgdsnlyasstmdyfssglmtekls 300
QY 301 QKNNLIFAVTENVNLYQNSELIPGTVGVLMSMDSSNVNQLIVDAYKIRSKVELEVR 360

Db 301 qknnlifavtenvnlyqyseleptgttvgvlsndsnvlqlivdaykirkvkelevr 360
QY 361 DLEPEEUSLSFWATCLNNENVPIGLKSCMGMLKIGDVWSSTEAKYRGCPQEKEKSTIKVG 420
Db 361 dlepeelsfratclnnenvipgkscmgmlkigdvsfreakyrgcpqekekstikvg 420
QY 421 FRSLSLIVQVTFDCDCACQQAEPNSHRCNCNGATFEGYCRGGWLSOCESSEEDYRP 480
Db 421 fkdslivqvtfdcdacqaaepnshrcncngatfegycrcgwgwlisqecseedyrp 480
QY 481 SQDQECSPRECGQPVSSORGECIGGCYCECIDSFRKITYGKCECIDSFRKITYGKCECIDSFRKITYG 540
Db 481 sqqdecspreqgqvcsrqgclcqvcvhsdflgkqyrcedafscrykqemcsq 540
QY 541 QCSGCDLCLSDWTGTYCNCTTRIDTCMSSNGLCSGRGKCECCSCVCIQPGSYGDRER 600
Db 541 qcsqgclclsdwtgtycnctrtdcmssnlgcsgqkceqscv:qpsqydrctek 600
QY 601 CPTCPDACTKKECKKFDRGALHDENTCNRYCRDTEESVKLKDNGDAVNCTYNE 660
Db 601 cptcpdaactfkkecveckkfarepymtentcnrycdelesvkelkdgkdvntcyne 660
QY 661 DDCVYREQYEDSSKSLIYVVEEPCPKGDPILWLSVGMALLIISALIWKLLT 720
Db 661 ddcvvrftqyvedssksliyvveepccpkqgdpilwlvslvsgmaliqjialliwkllt 720
QY 721 IHDREBFKAEEERAKRWDTANNPLYKEATSTPNTNIYRGT 762
Db 721 ihdrefakfeerarakrwdtannplykeatstpntrnyrgt 762

RESULT 10
ID AAY92444 standard; Protein; 788 AA.
XX
AC AAY92444;
DT 10-AUG-2000 { first entry}
XX
DE Variant human GPIIIa, alloantigen PLA2.
XX
GPIIIa; GPIIB; neurological disease; genotype; phenotype; diagnosis;
alloantigen; platelet antigen; PIA1; PLA2; Alzheimer's disease; stroke;
neurofibromatosis; Huntington's disease; depression; Parkinson's disease;
amyotrophic lateral sclerosis; multiple sclerosis; dementia.
XX
OS Homo sapiens.

XX
Key Location/Qualifiers
FT 1..26
FT /label= signal_peptide
FT Protein 27..788
FT /label= mature_protein
FT Msc-difference 33
FT /label= L33P
FT /note= "mutation caused by T19C substitution in
FT the coding sequence"
XX
PN WO200020634-A1.
XX
PD 13-APR-2000.
XX
PF 01-OCT-1999; 99WO-IB01696.
XX
PR 01-OCT-1998; 98US-0102624.
XX
PA (NOVA-) NOVA MOLECULAR INC.
XX
PI Schappert K;
XX
DR WPI; 2000-303801/26.
N-PSDB; AAA09363.

XX
PT Identifying a subject at risk for a neurological disease comprises
PT determination of genotype or phenotype of GPIIa or GPIb locus and
PT determining presence of variant GPIIa or GPIb allele or isoform
XX
PS Claim 16; Fig 4; 55pp; English.
XX
CC Two different forms of GPIIa, alloantigens PLA1 and PLA2 (for Platelet
CC Antigen 1 and 2) have been described and can be distinguished using a
monoclonal antibody. The rarer form, PLA2 has sustained a point mutation
CC at base 192 that causes a nucleotide change from a T to a C and thus a
leucine to proline amino acid substitution at residue position 33.
CC Identifying a subject at risk for a neurological disease comprises
CC determining the genotype or phenotype of the GPIIa or GPIb locus of a
subject, and determining the presence of a variant GPIIa or GPIb
allele or isoform, where the presence indicates an increased risk of
neurological disease. The method is useful for identifying a risk,
diagnosing and treating a neurological disease comprising Alzheimer's
disease, neurofibromatosis, Huntington's disease, depression, Parkinson's
disease, amyotrophic lateral sclerosis, multiple sclerosis, stroke, Parkinson's
disease and multi-infarct dementia.

SO Sequence 788 AA;

Query Match 99.0%; Score 4112; DB 21; Length 788;
Best Local Similarity 99.2%; Pred. No. 2e-282; Mismatches 0; Indels 0; Gaps 0;

OY 1 GPNICTTRGVSSCQCLAVSPMCAMCSDEALPLGSPPRCDEKLENLKDNCAPESIEFPVSE 60
Db 27 gpnicttrgvsscqclavspmcawcsdealplgspprcdekenlkdcapiesiefpvse 86
OY 61 ARVLEDRPLSDKGSGDSSQVTVQSPORIALRLRPPDSKNSIQVQFEDVYDPIVYMLD 120
Db 87 arvledrplsdkgsgdssqvqspqspqairlrpdsksnsiqvqfedvypdyiymld 146
OY 121 SYSMKDLWSTQNLGKRLATMRKLTNSNLRIGFAGFVDKPVSPWMVISPEALENPYCDM 180
Db 147 sysmkdlwlswtqnlglktqmkrltsnlrligfagfvdkpvspwmvispealenpycdm 206
OY 181 KTCIICLPGFGYHVLTLTDDOVRFNEEVKVKSVSRIRDAAPGGEDAMQAVYDCEKIGWRN 240
Db 207 ktclplmtgkyhvittdqvrfrfneevkkavsvrhrdapeggfdaimqatvcdekgwrn 266
OY 241 DASHLVLVFTDAKTHIALDGRLAGIVQPNDCQCHGSDNHYSASITMDYPSLGMTEKLS 300
Db 267 dashlvlvftdakthialdgrlagivqpnqgqhgqsdnlyasitmdyplgimtekls 326
OY 301 ORNINLTFAVTENVVNLXQNSELPGLPTTVGVLMSDSSNVQLIDAYGKRSKVELEVR 360
Db 327 qknniifavkenvnlyqnselplgpttvgvlsdnsvnqilividaygkirkvelevr 386
OY 361 DIFPEELSLFSNATCLUNNEVIGLKSQMGKIGDTSFSIAKVRQCPQEEKSFVTPVG 420
Db 387 dpeelslfsnafatclunnevlgksqmgkigdtsfsiakvrcpgeeksfvtpvg 446
OY 421 FKDSLIVQVTFDCACQAQAPNSHRCNNNGTFRGVRCRGPGWLGSOCECSEDYRP 480
Db 447 fksdlivqvtfddccacqaqgepnsnrcngngtfcgvcrcgpgwlgsgcscseedyrp 505
OY 481 SQQDECSPREQGPPVCSQRGECIICLGGQCVCHSSDFGKTFGKVCCECDPSCVRYKGEMCSGHG 540
Db 507 seqdecspreqgppvcsqrgecigggcvchssdfgkftgkyceddfscvrykgemcsghg 566
OY 541 QCCGGDCLCPSWPTGVCNCNTTRDPCMSUNGLLGSGRKCECGSCVCTPGSGDTEK 600
Db 567 qccggdclcdsawtggcvcncttrtdtcmssqglcsgrgkcecgscvcigpgsydtcek 626
OY 601 CPTCPDACTFKCEVCEKKFDRGALDENTCNRYCDEIESVKEELKDTGDAVNTYKE 660
Db 627 cptcpactfkcevcekkfdrpymentcnrycdeiesvkeelkdtgdaavntyke 686

QY 661 DDCWVRFQYESSGKILYVVEBPCPKGPILWILVLSGAGILGILMLWILIT 720
CC ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 687 ddcwvrfqyessgkilslyvvebpcpkgpilwilsngailglalilwllit 746
QY 721 IHRKEFAKEERARAKWDTANNPLKEASSTETNTYRST 762
CC ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 747 lhrkefakfeerarakwdtannplkeasstetntyrgt 788
QY RESULT 11
Db AAV92452
ID AAV92452 standard; Protein; 762 AA.
AC AAV92452;
XX
DT 10-AUG-2000 (first entry)
XX
DE GPIIa variant Arg214TrP.
XX
KW GPIIa; GPIIb; neurological disease; genotype; phenotype; diagnosis;
KW alloantigen; platelet antigen; PLA1; PLA2; Alzheimer's disease; stroke;
KW neurofibromatosis; Huntington's disease; depression; Parkinson's disease;
KW amyotrophic lateral sclerosis; multiple sclerosis; dementia.
XX
OS Homo sapiens.
XX
PN WO20020634-A1.
XX
PD 13-APR-2000.
XX
PP 01-OCT-1999; 99W0-1B01696.
XX
PR 01-OCT-1998; 980S-0102624.
XX
PA (NOVA-) NOVA MOLECULAR INC.
XX
PI Schaper K;
XX
DR 2000-303801/26.
XX
PT Identifying a subject at risk for a neurological disease comprises
PT determining of genotype or phenotype of GPIIa or GPIb locus and
PT determining presence of variant GPIIa or GPIb allele or isoform
XX
PS Disclosure; Page -; 55pp; English.
XX
Known polymorphisms in GPIIa that may be determined to be variants
CC using the methods of the invention include AAY2447-56, which are variant
GPIIa sequences derived from the mature protein.
CC Identifying a subject at risk for a neurological disease comprises
CC determining the genotype or phenotype of the GPIIa or GPIb locus of a
subject, and determining the presence of a variant GPIIa or GPIb
allele or isoform, where the presence indicates an increased risk of
neurological disease. The method is useful for identifying Alzheimer's
CC diagnosing and treating a neurological disease comprising Alzheimer's
disease, neurofibromatosis, Huntington's disease, depression,
CC amyotrophic lateral sclerosis, multiple sclerosis, stroke, Parkinson's
disease and multi-infarct dementia.
CC Note: This sequence is not given in the specification it was created
CC from the wild type human GPIIa sequence which appears in Figure 3.
XX
SO Sequence 762 AA;

Query Match 99.0%; Score 4111; DB 21; Length 762;
Best Local Similarity 99.2%; Pred. No. 2.3e-282; Mismatches 0; Indels 0; Gaps 0;

OY 1 GPNICTTRGVSSCQCLAVSPMCAMCSDEALPLGSPPRCDEKLENLKDNCAPESIEFPVSE 60
Db 1 gpnicttrgvsscqclavspmcawcsdealplgspprcdekenlkdcapiesiefpvse 60
OY 61 ARVLEDRPLSDKGSGDSSQVTVQSPORIALRLRPPDSKNSIQVQFEDVYDPIVYMLD 120

PA (NOVA-) NOVA MOLECULAR INC.
 XX
 PT Schappert K;
 XX DR WI; 2000-303801/26.

PT Identifying a subject at risk for a neurological disease comprises determination of genotype or phenotype of GPIIa or GPIb locus and determining presence of variant GPIIa or GPIb allele or isoform.

PS Disclosure: Page -; 55pp; English.

XX Known polymorphisms in GPIIa that may be determined to be variants using the methods of the invention include AAY92447-56, which are variant GPIIa sequences derived from the mature protein.

CC Identifying a subject at risk for a neurological disease comprises determining the genotype or phenotype of the GPIIa or GPIb locus of a subject, and determining the presence of a variant GPIIa or GPIb allele or isoform, where the presence indicates an increased risk of neurological disease. The method is useful for identifying a risk, diagnosis and treating a neurological disease comprising Alzheimer's disease, neurofibromatosis, Huntington's disease, depression, anoxirotrophic lateral sclerosis, multiple sclerosis, stroke, Parkinson's disease and multi-infarct dementia.

CC Note: This sequence is not given in the specification; it was created from the wild type human GPIIa sequence which appears in Figure 3.

CC Sequence 762 AA;

Query Match 99.0%; Score 4111; DB 21; Length 762; Best Local Similarity 99.2%; Pred. No. 2.3e-282; Matches 756; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 GPNICTRGSSCQCLAVSPMCAMCSDAELPLGSPRCDLKENLKDCAPESTIEPPSE 60
 Db 1 gpnictrgsscqcavspmcawcsdealplgsprcdlkenlkdcapestieppse 60

QY 61 ARVLEDRLPLSDKGSDSDSQVQSPQRALRDRDDSNFSIQVRQVDYPVDIYMDL 120
 Db 61 arvledrlplsdkgssqvtqspqrarlrpdsknfsiqvrqvdypvdiyimdl 120

QY 121 SYSMKDLSNQNTGKLATOMKLTNSRIGAFVRKPVSEMYISPPALENPYCDM 180
 Db 121 sysmkdlsnqntgklatqmrktsnrlqgafvdkpvstmyisppalenpycdm 180

QY 181 KTTCLPMFGKHKVLTQDVTRENEEVKQSVSNRDAPEGGDAMQATVDEKIGWRN 240
 Db 181 kttclpmfgkhwltqdtqtrneevkqsvsnrdapeggdaimatvdekekgwrn 240

QY 241 DASHLIVFTDAKHALDORLAGTVQPDQCHVGNSHISASTMDYPSLGLMTEKLS 300
 Db 241 dashlifftdakthaldrlagivqpdqchvgnsahyastmdyplsglmtekls 300

QY 301 QKNTILFAVENVNLYQNSLELPGTQVGVLSMDSNVLOIIVDAYGKRSVLEVR 360
 Db 301 qknntilfaatenvnlyqnslelpqtvgvlsmdsnvloivdaygkrsvevr 360

QY 361 DIPPELSLSFNATCLNEVTPGKSCMGLKIGTVPSIEAKYRGCPQEKSFTKPGV 420
 Db 361 dipeelsisfnatclnevtpgkscmglkigtvpsieakvrgcprqeksfktpgv 420

QY 421 FKSLSIVQVTFDCCACACQAQAEPRNSHRQNNGNTFECGVCRCPGWLGSQCCESEEDYRP 480
 Db 421 fkslsivqvtfdccacacqaqaeprnshrqnnngntfecgvcrccpgwlgssqcceedyrp 480

QY 481 SQDQECSPREGQPVCSQRCGKSCMGLKIGTVPSIEAKYRGCPQEKSFTKPGV 540
 Db 481 sqdqcspregqpvcsqrcgkscmglkigtvpsieakvrgcprqeksfktpgv 540

QY 541 QCSGGDCCLSDTGYVCNCNTARTDTCMSNGNLCSGRGKCGGCVCIQPGSYDTEK 600
 Db 541 qcsqcgcacldsdawtggyencttrtdtcmnsngllcsgrgkcecgcvciqpgsygtcek 600

QY 601 CPGCPDACKFKECKVCEKKFDPRGALHENTCNCYRCDEIESYKELDTGKRAVNCTYNE 660
 Db 601 cpgcpdactfkkeckvcekkfdprgalmhentcncycdeiesykeldtgkavcnctyne 660

QY 661 DDCWVRFQYEDSSGKSLIVVVEPECPKGDPDILVWLSVMAAILGLAALIWKLT 720
 Db 661 ddcwvrfqyedssgkslivvvepecpkgdpdilvwlsvmaailglalaaliklt 720

QY 721 IHRKEFAKEFEEARAKWDTANNPLKYEATSTFTNITYRG 762
 Db 721 ihorkefakekerearakwtdannplkyeaustftnityrg 762

RESULT 12

AAV92454
 ID AAV92454 standard; Protein; 762 AA.
 XX AC AAV92454;
 XX DT 10-AUG-2000 (first entry)
 XX DE GPIIa variant Pro407Ala.
 XX
 KW GPIIa; GPIb; neurological disease; genotype; phenotype; diagnosis; alloantigen; platelet antigen; PIA1; PIA2; Alzheimers disease; stroke; neurofibromatosis; Huntingtons disease; depression; Parkinsons disease; amyotrophic lateral sclerosis; multiple sclerosis; dementia.
 KW Homo sapiens.
 XX PN W020020634-A1.
 XX PD 13-APR-2000.
 XX PF 01-OCT-1999; 99WO-1B01696.
 XX PR 01-OCT-1998; 98US-0102624.
 XX

Db 541 qcsqcgcacldsdawtggyencttrtdtcmnsngllcsgrgkcecgcvciqpgsygtcek 600

QY 601 CPTCDACTFKKCEVECKKEDRGALHDENTCNRYCDELESKELKDTGKDAVNCTYKNE 660
 ||||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 601 cptcpdactfkkeveckkfrepymtentcnrycdeleskkelkdtgkdavnctyke 660
 QY 661 DDCVRFQYEDSGSKSILVYVREPECPCKGPDIVLILSVMGATLIGLAALIWKLLIT 720
 ||||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 661 ddcvrfqyedsgksilvverepcpcgpdilvllsvalgalligalaalikllit 720
 QY 721 IHDRKEFAKEEERAKWDTANNPYKEATSTFNITYRGT 762
 ||||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 721 ihdrkefakfeerarakwtannplykeatstfnityrgt 762
 QY RESULT 13
 AAY9455
 ID AAY9455 standard; Protein; 762 AA.
 XX
 AC AAY9455;
 XX
 DT 10-AUG-2000 (first entry)
 DE GPIIa variant Arg636Cys.
 XX
 KW GPIIa; GPIb; neurological disease; genotype; phenotype; diagnosis;
 KW allantigen; platelet antigen; PLAI; PLAI2; Alzheimers disease; stroke;
 KW neurofibromatosis; Huntingtons disease; depression; Parkinsons disease;
 KW amyotrophic lateral sclerosis; multiple sclerosis; dementia.
 OS Homo sapiens.
 XX
 WO200020634-A1.
 PN
 XX
 PD 13-APR-2000.
 XX
 PF 01-OCT-1999; 99WO-1B01696.
 PR 01-OCT-1999; 98US-0102624.
 XX
 PA (NOVA-) NOVA MOLECULAR INC.
 XX
 PI Schappert K;
 XX
 DR WPI; 2000-303801/26.
 XX
 PT identifying a subject at risk for a neurological disease comprises
 PT determining presence of variant GPIIa or GPIb allele or isoform
 XX
 Disclosure; Page - ; 55pp; English.
 XX
 CC known polymorphisms in GPIIa that may be determined to be variants
 CC using the methods of the invention include Aay92447-56, which are variant
 CC GPIIa sequences derived from the mature protein.
 CC identifying a subject at risk for a neurological disease comprises
 CC determining the genotype or phenotype of the GPIIa or GPIb locus of a
 CC subject, and determining the presence of a variant GPIIa or GPIb
 CC allele or isoform, where the presence indicates an increased risk of
 CC neurological disease. The method is useful for identifying a risk,
 CC diagnosing and treating a neurological disease comprising Alzheimer's
 CC disease, neurofibromatosis, Huntington's disease, depression,
 CC amyotrophic lateral sclerosis, multiple sclerosis, stroke, Parkinson's
 CC disease and multi-infarct dementia.
 CC Note: This sequence is not given in the specification it was created
 XX from the wild type human GPIIa sequence which appears in Figure 3.
 SQ Sequence 762 AA;

QY RESULT 14
 AAY92449
 ID AAY92449 standard; Protein; 762 AA.
 XX
 AC AAY92449;
 XX
 DT 10-AUG-2000 (first entry)
 XX
 DE GPIIa variant Asp18Tyr.
 XX
 KW GPIIa; GPIb; neurological disease; genotype; phenotype; diagnosis;
 KW allantigen; platelet antigen; PLAI; PLAI2; Alzheimers disease; stroke;
 KW neurofibromatosis; Huntingtons disease; depression; Parkinsons disease;
 KW amyotrophic lateral sclerosis; multiple sclerosis; dementia.
 XX
 OS Homo sapiens.
 XX
 WO200020634-A1.
 XX
 PD 13-APR-2000.

Query Match 99.0%; Score 4111; DB 21; Length 762;
 Best Local Similarity 99.2%; Pred. No. 2.3e-282;
 Matches 756; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

XX 01-OCT-1999; 99WO-IB01696.
 PF XX 01-OCT-1998; 98US-0102624.
 PR XX (NOVA-) NOVA MOLECULAR INC.
 PA XX PI Schappert K;
 DR XX WPI; 2000-303801/26.

XX Identifying a subject at risk for a neurological disease comprises determination of genotype or phenotype of GPIIIa or GPIb locus and determining presence of variant GPIIIa or GPIb allele or isoform

XX Disclosure: Page -; 55pp; English.

CC Known polymorphisms in GPIIIa that may be determined to be variants using the methods of the invention include AAY92447-56, which are variant GPIIIa sequences derived from the mature protein.

CC Identifying a subject at risk for a neurological disease comprises determining the genotype or phenotype of the GPIIIa or GPIb locus of a subject, and determining the presence of a variant GPIIIa or GPIb allele or isoform, where the presence indicates an increased risk of neurological disease. The method is useful for identifying a risk, diagnosing and treating a neurological disease comprising Alzheimer's disease, neurofibromatosis, Huntington's disease, depression, amyotrophic lateral sclerosis, multiple sclerosis, stroke, Parkinson's disease and multi-infarct dementia.

CC Note: This sequence is not given in the specification it was created from the wild type human GPIIIa sequence which appears in Figure 3.

XX Sequence 762 AA;

Query Match 98.9%; Score 4110; DB 21; Length 762;
 Best Local Similarity 99.2%; Pred. No. 2.7e-28; Matches 756; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 GPMICTRGVSSCQCLAVSPMCAWSDEAULPLGSPRDLKENDNCAPESIEPVSE
 Db 1 gpictctrgrgssccqclavspmcawsdeaplgsprdlkenlkdcapessiefpse. 60

QY 61 ARYLEDLRPLDKSGDSSQVTOVSPORIALRPPDKSKNFESIVROVEDYVDTYLMQ
 Db 61 arvleplsdksqgssqvqspqrrialrpdsknfisqvrvdyvbdyyimy. 120

QY 121 SYSMKDDIWSLQNLGKTLATRKRLTSNLRGFGAIVDKPKSPYWISPPALEPCYDM
 Db 121 sysmkddiwslnlgktakomrktsnlrigfawdkpyspymisppalepcydm. 180

QY 181 KITCLPMEFGIKVHLTDQVRFNEEVKKOSVRNRDAPEGGFDATMQATVODEKIGWRN
 Db 181 ktclcpmefgikvltdqvtfnneevkkqksrsndapegfdaimgatvodekigwrn. 240

QY 241 DASHLYFTTDKTHALDGLAGTVQPNDOCHVGSNDHYSASTMDYPSLGMTEKLS
 Db 241 dashlyfttdkthaldglagivqpnndqchvgsndhysastmdyplgmteks. 300

QY 301 QRNINLIFAVTENVWYKONSELIGCTVYLSMDSNNVYQLIDAYGIRSKVELEVR
 Db 301 qrninlifavtenvlyqnselipgqvlvdsnnvylqldaygirkvelevr. 360

QY 361 DPEEULSFLFATCLNNEVTPGKLSOMGLKIGDTWSFSLRKAQVRCPOENEKSFTKPGV
 Db 361 dpeeulsfatclnnenvlyqnselipgqvlvdsnnvylqldaygirkvelevr. 420

QY 421 FDSLIVQVTFTDCDCACQAOEPNSHRCNNNGTFCGCGRCGPGWLGSOCCESEDYRP
 Db 421 fdlivqvtfdccacqapepnsnrcnngtfcgvcrcgpgwlgsocceseyrp. 480

QY 481 SQDDESPREOPVOSORGECLGQCVCHSSDFGKIGKCECCDFSCVWKGEMCSHG
 XX Sequence 762 AA;

Db 481 sqddecspregqpvcsqrgceclcgqvcvhsdfgkitgkyrceduts:cyrykgemcsqg 540

QY 541 QCSGQICLDDWDTGVCNC'TRITCMSSNGLCMSGKCECSCVCQPGSYDTEK 600

Db 541 qcsqdclcdsdwtgqyancntrtdtcmssnglicsgrgjckcecgscv:qpgsydtek 600

QY 601 CPTCPDACTFKKECYCKKEDRGALHDENTCNRYCRDDELSVKELKDQKDAVNTYKNE 660

Db 601 cptcpdactfkkecyckkfrerepymtentcnrycrdelsvkeldqkdaavntyke 660

QY 661 DDCVVRROYYEDSSGSKSILYVVEERBCKPKGPDILVVLSLSGMAGLILIGAALLWKLIT 720

Db 661 ddcvififqyedssgksilywveepccpkqpdaiwvlsivmgailiaalliwkliit 720

QY 721 IHDRIKFKAFKEERAKWATNANNLYKRASTTINITYRGT 762

Db 721 hdkrfakfeerarakwattanpyleatsttntyrgt 762

RESULT 15
 AAY92453
 ID AAY92453 standard; Protein: 762 AA.
 XX AC AAY92453;
 PR 10-ADG-2000 (first entry)
 XX DE GPIIIa variant Cys374Tyr.
 XX KW GPIIIa; GPIIIb; neurological disease; genotype; phenotype; diagnosis; alloantigen; platelet antigen; PI1; PI2; Alzheimer's disease; stroke; neurofibromatosis; Huntington's disease; depression; Parkinson's disease; amyotrophic lateral sclerosis; multiple sclerosis; dementia.
 XX OS Homo sapiens.
 XX PN WO200020634-A1.
 XX PD 13-APR-2000.
 PR XX 01-OCT-1999; 99WO-IB01696.
 PA XX (NOVA-) NOVA MOLECULAR INC.
 PI Schappert K;
 DR XX WPI; 2000-303801/26.

XX Identifying a subject at risk for a neurological disease comprises determining of genotype or phenotype of GPIIIa or GPIb locus and determining presence of variant GPIIIa or GPIb allele or isoform

XX Disclosure: Page -; 55pp; English.

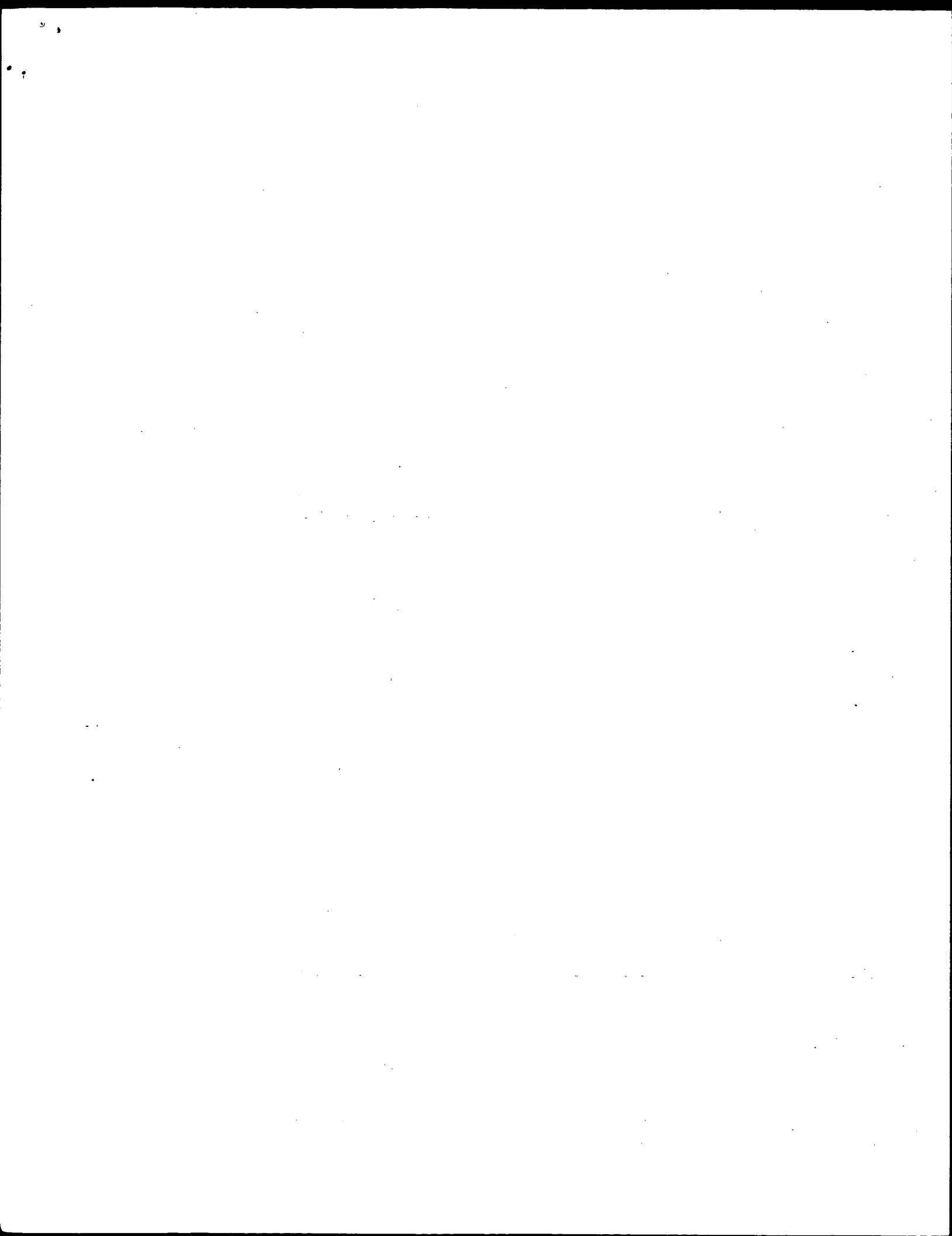
CC Known polymorphisms in GPIIIa that may be determined to be variants using the methods of the invention include AAY92447-56, which are variant GPIIIa sequences derived from the mature protein.

CC Identifying a subject at risk for a neurological disease comprises determining the genotype or phenotype of the GPIIIa or GPIb locus of a subject, and determining the presence of a variant GPIIIa or GPIb allele or isoform, where the presence indicates an increased risk of neurological disease. The method is useful for identifying Alzheimer's disease, diagnosing and treating a neurological disease comprising Alzheimer's disease, neurofibromatosis, Huntington's disease, depression, amyotrophic lateral sclerosis, multiple sclerosis, stroke, Parkinson's disease and multi-infarct dementia.

CC Note: This sequence is not given in the specification it was created from the wild type human GPIIIa sequence which appears in Figure 3.

XX Sequence 762 AA;

Search completed: May 19, 2002, 10:38:56
Job time: 9118 sec



Copyright (c) 1993 - 2000 Compugen Ltd.

Om protein - protein search, using sw model

Run on: May 19, 2002, 08:07:51 ; Search time 36.84 Seconds

Perfect score: US-09-673-302A-1

Sequence: 1 GPNICTRGVSSCOCCLAVS.....NPNPLYKEATSTFTNITYRGT 762

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 2425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA; *
1: /cgn2_6/prodata/2/1aa/5A_COMB.pep: *
2: /cgn2_6/prodata/2/1aa/5B_COMB.pep: *
3: /cgn2_6/prodata/2/1aa/6A_COMB.pep: *
4: /cgn2_6/prodata/2/1aa/6B_COMB.pep: *
5: /cgn2_6/prodata/2/1aa/pctm5_COMB.pep: *
6: /cgn2_6/prodata/2/1aa/backfile1.pep: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	4119	99.2	788	2 US-07-228-215-32
2	4119	99.2	788	4 US-08-938-085A-2
3	3768	90.7	718	1 US-08-444-792-4
4	3768	90.7	718	1 US-08-445-042-4
5	2339	56.3	799	1 US-08-054-07C-2
6	2156	51.9	788	2 US-07-228-215-27
7	2156	51.9	788	4 US-08-938-085A-27
8	1831	44.1	798	2 US-07-228-215-30
9	1831	44.1	798	4 US-08-938-085A-30
10	1676	40.3	577	2 US-07-728-215-29
11	1676	40.3	577	4 US-08-938-085A-29
12	1535.5	37.0	846	2 US-07-728-215-33
13	1535.5	37.0	846	4 US-08-938-085A-33
14	1513.5	36.4	769	2 US-08-789-078-1
15	1513.5	36.4	769	2 US-08-789-078-1
16	1513.5	36.4	769	2 US-08-72-062A-45
17	1513.5	36.4	769	2 US-07-72-015-31
18	1513.5	36.4	769	4 US-08-938-085A-31
19	1513.5	36.4	769	5 PCT-US5-01886-1
20	1513.5	36.4	769	5 PCT-US6-01314-45
21	1433.5	32.3	676	3 US-08-930-172-10
22	1335.2	32.3	676	4 US-09-975-119-10
23	1335.2	32.3	676	5 PCT-US6-01314-57
24	1331.2	32.0	768	1 US-08-554-455-4
25	1336.5	31.9	769	1 US-08-554-455-6
26	1278.0	30.8	240	5 PCT-0444-6
27	1278.0	30.8	266	2 US-07-008-457-3

ALIGNMENTS

RESULT	SEQUENCE	SEQUENCE	OPERATING SYSTEM	SOFTWARE	PATENT	RELEASE	NUMBER	VERSION	DATE
US-07-728-215-32	Sequence 32, Appl	Sequence 32, Appl	PC-DOS/MS-DOS	PatentIn	Sheppard, Dean	5962443	1	1.0	US/07728215
	Sequence 32, Appl	Sequence 32, Appl			Quaranta, Vito				
	Sequence 4, Appl	Sequence 4, Appl			Pytelia, Robert				
	Sequence 2, Appl	Sequence 2, Appl			TITLE OF INVENTION: A No. 5962643el Integrin Beta				
	Sequence 27, Appl	Sequence 27, Appl			TITLE OF INVENTION: Thereof				
	Sequence 30, Appl	Sequence 30, Appl			NUMBER OF SEQUENCES: 43				
	Sequence 30, Appl	Sequence 30, Appl			CORRESPONDENCE ADDRESS:				
	Sequence 29, Appl	Sequence 29, Appl			ADDRESSEE: Pretty, Schroeder, Brueggemann & Cl				
	Sequence 29, Appl	Sequence 29, Appl			STREET: 4370 La Jolla Village Drive, Suite 700				
	Sequence 33, Appl	Sequence 33, Appl			CITY: San Diego				
	Sequence 33, Appl	Sequence 33, Appl			STATE: California				
	Sequence 33, Appl	Sequence 33, Appl			COUNTRY: United States of America				
	Sequence 33, Appl	Sequence 33, Appl			ZIP: 92122				
	Sequence 33, Appl	Sequence 33, Appl			COMPUTER READABLE FORM:				
	Sequence 33, Appl	Sequence 33, Appl			MEDIUM TYPE: FLOPPY disk				
	Sequence 33, Appl	Sequence 33, Appl			COMPUTER: IBM PC compatible				
	Sequence 33, Appl	Sequence 33, Appl			OPERATING SYSTEM: PC-DOS/MS-DOS				
	Sequence 33, Appl	Sequence 33, Appl			SOFTWARE: PatentIn Release #1.0, Version #1.25				
	Sequence 33, Appl	Sequence 33, Appl			CURRENT APPLICATION DATA:				
	Sequence 33, Appl	Sequence 33, Appl			APPLICATION NUMBER: US/07728,215				
	Sequence 33, Appl	Sequence 33, Appl			FILING DATE: 19910711				
	Sequence 33, Appl	Sequence 33, Appl			CLASSIFICATION: 435				
	Sequence 33, Appl	Sequence 33, Appl			ATTORNEY/AGENT INFORMATION:				
	Sequence 33, Appl	Sequence 33, Appl			NAME: Campbell, Catnlyn A.				
	Sequence 33, Appl	Sequence 33, Appl			REGISTRATION NUMBER: 31,815				
	Sequence 33, Appl	Sequence 33, Appl			REFERENCE/DOCKET NUMBER: P31 8717				
	Sequence 33, Appl	Sequence 33, Appl			TELECOMMUNICATION INFORMATION:				
	Sequence 33, Appl	Sequence 33, Appl			TELEPHONE: (619) 535-9001				
	Sequence 33, Appl	Sequence 33, Appl			FAX: (619) 535-9949				
	Sequence 33, Appl	Sequence 33, Appl			INFORMATION FOR SEQ ID NO: 32:				
	Sequence 33, Appl	Sequence 33, Appl			SEQUENCE CHARACTERISTICS:				
	Sequence 33, Appl	Sequence 33, Appl			LENGTH: 788 amino acids				
	Sequence 33, Appl	Sequence 33, Appl			TYPE: AMINO ACID				
	Sequence 33, Appl	Sequence 33, Appl			STRANDBENESS: single				
	Sequence 33, Appl	Sequence 33, Appl			TOPOLOGY: linear				
	Sequence 33, Appl	Sequence 33, Appl			US-07-728-215-32				
	Sequence 33, Appl	Sequence 33, Appl			Query Match	99.2%	Score	4119	DB 2;
	Sequence 33, Appl	Sequence 33, Appl			BEST LOCAL SIMILARITY	99.3%	PRED. NO.	0;	MISMATCHES
	Sequence 33, Appl	Sequence 33, Appl			MATCHES	757			5;
	Sequence 33, Appl	Sequence 33, Appl			QY	1	1		
	Sequence 33, Appl	Sequence 33, Appl			DB	27	1		
	Sequence 33, Appl	Sequence 33, Appl			2	27	1		

CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/938,085A
 FILING DATE: 26-SEP-1997
 CLASSIFICATION: 435
 PRIORITY APPLICATION DATA:
 APPLICATION NUMBER: US 07/728,215
 FILING DATE: 11-JUL-1991
 ATTORNEY/AGENT INFORMATION:
 NAME: Parent, Annette S.
 REGISTRATION NUMBER: 42,058
 REFERENCE/DOCKET NUMBER: 023070-080210US
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (415) 576-0200
 TELEFAX: (415) 576-0300
 INFORMATION FOR SEQ ID NO: 32:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 788 amino acids
 TOPOLOGY: linear
 STRANDEDNESS:
 MOLECULE TYPE: protein
 ; US - 08 - 938 - 085A - 32

Query Match 99.2%; Score 4119; DB 4; Length 788;
 Best Local Similarity 99.3%; Pred. No. 0; Gaps 0;
 Matches 757; Conservative 0; Mismatches 5; Indels 0;

QY 1 GENICTTRGVSSCQCLAVSPMCACWSDDEALPLGSPRCDLKENLKDNCAPESTEFPVSE 60
 Db 27 GNICITTRGVSSCQCLAVSPMCACWSDDEALPLGSPRCDLKENLKDNCAPESTEFPVSE 86
 QY 61 ARVLEDPLSDKGSGDSSQVITQSPORIALRRLPDDSKNSIQVROVEDYPVDIYMLD 120
 Db 87 ARVLEDPLSDKGSGDSSQVITQSPORIALRRLPDDSKNSIQVROVEDYPVDIYMLD 146
 QY 121 SYSMKDDLWSIONLGTRKLTOMRKLTNSNLRIGGAFVSPWYISPEALENPYCM 180
 Db 147 SYSMKDDLWSIONLGTRKLTOMRKLTNSNLRIGGAFVSPWYISPEALENPYCM 206
 QY 181 KITCLPMFGYKHVLITDQVTRFNEVKKOSVSRNRDAPEGFDAIMQATVCDEKIGRN 240
 Db 207 KITCLPMFGYKHVLITDQVTRFNEVKKOSVSRNRDAPEGFDAIMQATVCDEKIGRN 266
 QY 241 DASHLVLVFTDAKTHIALDGRLAGITVQPNQGCHVGSNDHYSASTMDYPSLGMTEKUS 300
 Db 267 DASHLVLVFTDAKTHIALDGRLAGITVQPNQGCHVGSNDHYSASTMDYPSLGMTEKUS 326
 QY 301 ORNINILFAVENVVNLVLYONSELIPGTTVGLSMDSSNVNQLIVDAYGKRSVLEVR 360
 Db 327 ORNINILFAVENVVNLVLYONSELIPGTTVGLSMDSSNVNQLIVDAYGKRSVLEVR 386
 QY 361 DPEELSLSFSNATCLNEVPIGLKSCMGKIGDVSFSTEAKVRCPOEKEKPTIKVG 420
 Db 387 DPEELSLSFSNATCLNEVPIGLKSCMGKIGDVSFSTEAKVRCPOEKEKPTIKVG 446
 QY 421 FDLSLQVWIFDCDCACQAOAEPNSHRCNNNGFPECGVRCGGWLGSOCESEEDYRP 480
 Db 421 FDLSLQVWIFDCDCACQAOAEPNSHRCNNNGFPECGVRCGGWLGSOCESEEDYRP 480
 QY 447 FDLSLQVWIFDCDCACQAOAEPNSHRCNNNGFPECGVRCGGWLGSOCESEEDYRP 506
 Db 447 FDLSLQVWIFDCDCACQAOAEPNSHRCNNNGFPECGVRCGGWLGSOCESEEDYRP 506
 QY 507 SQDQECSPREGQPVCSQRGECLGCGCVCVHSDFGKIGKYCECDFSCVRYKEMC5GH 566
 Db 507 SQDQECSPREGQPVCSQRGECLGCGCVCVHSDFGKIGKYCECDFSCVRYKEMC5GH 566
 QY 541 QSCGCLCSDWGGYCNCTTRDTCMSMGLCSCGKCEGSCVCIQPESYGDRCEK 600
 Db 541 QSCGCLCSDWGGYCNCTTRDTCMSMGLCSCGKCEGSCVCIQPESYGDRCEK 626
 QY 567 QSCGCLCSDWGGYCNCTTRDTCMSMGLCSCGKCEGSCVCIQPESYGDRCEK 626
 Db 567 QSCGCLCSDWGGYCNCTTRDTCMSMGLCSCGKCEGSCVCIQPESYGDRCEK 626
 QY 601 CPTCPDACTKKECCKFDRGALHDENTCNRYCRDELESVKELKDGKDAVNCTYKNE 660
 Db 627 CPTCPDACTKKECCKFDRGALHDENTCNRYCRDELESVKELKDGKDAVNCTYKNE 686
 QY 661 DDCVVRFOYEDSSGKSLVYVVEPECKGPDILVLLSVMGAILLGLAALLIWKLIT 720
 Db 687 DDCVVRFOYEDSSGKSLVYVVEPECKGPDILVLLSVMGAILLGLAALLIWKLIT 746
 QY 721 IHDRKEFAKEERAKKDTANNPPLYKATSTFTNITYRT 762
 Db 747 IHDRKEFAKEERAKKDTANNPPLYKATSTFTNITYRT 788
 QY RESULT 2
 US-08-938-085A-32
 ; Sequence 32, Application US/08938085A
 ; Patent No. 6339148
 ; GENERAL INFORMATION:
 ; APPLICANT: Quaranta, Vito
 ; APPLICANT: Pytel, Robert
 ; TITLE OF INVENTION: A No. 6339148ei Integrin Beta Subunit and Uses
 ; TITLE OF INVENTION: Thereof
 ; NUMBER OF SEQUENCES: 62
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Townsend and Townsend and Crew LLP
 ; STREET: Two Embarcadero Center, Eighth Floor
 ; CITY: San Francisco
 ; STATE: California
 ; COUNTY: USA
 ; ZIP: 94111-3834
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30

RESULT 3
US-08-44-792-4
Sequence 4, Application US/08444792
; Patent No. 5726037
GENERAL INFORMATION:
APPLICANT: Bodary, Sarah C.
APPLICANT: Gorman, Cornelia M.
APPLICANT: McLean, John W.
APPLICANT: Naper, Mary A.
TITLE OF INVENTION: METHOD FOR PREPARING WATER SOLUBLE
TITLE OF INVENTION: POLYPEPTIDES
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/444,792
FILING DATE: 19-MAY-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/380227
FILING DATE: 30-JAN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/444,792
FILING DATE: 28-MAR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/921337
FILING DATE: 13-JAN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/444490
FILING DATE: 01-DEC-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/290224
FILING DATE: 22-DEC-1988
ATTORNEY/AGENT INFORMATION:
NAME: Lee, Wendy M.
REGISTRATION NUMBER: 00 000
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-1994
TELEFAX: 415/952-9881
TELEX: 910371-7168
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 718 amino acids
TYPE: Amino Acid
TOPOLGY: Linear
US-08-444-792-4

Query Match 90.7%; Score 3768; DB 1; Length 718;
Best Local Similarity 99.3%; Pred. No. 2.5e-205; Matches 687; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 GPNICTTGRGSSQOCLAVSPMKAWCSDAELPGLSPRDLKENLKLKNCAPESIEFPVSE 86
Db 687 DDCVVRQFYEDSSGSKSLIVVVEEPCPKGD 745
QY 721 IHRKERAKFEERAKRWDTANPYLEAKSTFTNTYRG 762
Db 747 IHRKERAKFEERAKRWDTANPYLEAKSTFTNTYRG 788

RESULT 4
US-08-445-042-4
Sequence 4, Application US/08445042
; Patent No. 5726290
GENERAL INFORMATION:
APPLICANT: Bodary, Sarah C.
APPLICANT: Gorman, Cornelia M.
APPLICANT: McLean, John W.
APPLICANT: Naper, Mary A.
TITLE OF INVENTION: METHOD FOR PREPARING WATER SOLUBLE
TITLE OF INVENTION: POLYPEPTIDES
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/445,042
FILING DATE: 01-DEC-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/380227
FILING DATE: 30-JAN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/445,042
FILING DATE: 28-MAR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/921337
FILING DATE: 13-JAN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/444490
FILING DATE: 01-DEC-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/290224
FILING DATE: 22-DEC-1988
ATTORNEY/AGENT INFORMATION:
NAME: Lee, Wendy M.
REGISTRATION NUMBER: 00 000
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-1994
TELEFAX: 415/952-9881
TELEX: 910371-7168
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 718 amino acids
TYPE: Amino Acid
TOPOLGY: Linear
US-08-445-042-4

Query Match 90.7%; Score 3768; DB 1; Length 718;
Best Local Similarity 99.3%; Pred. No. 2.5e-205; Matches 687; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 61 ARVLEDPLSDKGSDSSVTOVSPORALLRPPDSNSFSQLQRQVEDYPVDIYJMDL 120
Db 87 ARVLEDPLSDKGSDSSVTOVSPORALLRPPDSNSFSQLQRQVEDYPVDIYJMDL 146
QY 121 SYSMKDDLSNISTQGTLKLTOMRRLTSNLRIGGAFVDRPKVSPSMYSSPPEALENPYCDM 180
Db 147 SYSMKDDLSNISTQGTLKLTOMRRLTSNLRIGGAFVDRPKVSPSMYSSPPEALENPYCDM 206
Db 207 KTTCLPMFGYKHVLTILTDQYTRFNEEKKVSNSNRDAPGGDAMQNTVDEKIGRN 266
QY 241 DASHLUVFTDAKTHIALDGRLAGIVQPNQGCHGSDNHYASSTTMYPSLGMTELS 300
Db 327 QRNINLIFATENVNLYQNSLEIPGTVGVLMSDNVNLQIVDAGKRSKVELEVR 386
QY 361 DLPEEULSLSNATCLNEVPGLSCMGKIGDVSFSSTEAKVQSCPUERKEKFTIKVG 420
Db 387 DLPEEULSLSNATCLNEVPGLSCMGKIGDVSFSSTEAKVQSCPUERKEKFTIKVG 446
QY 421 FKDSLIVQVTFDCACQACQQAEPFSPHRCCNGNGFECGVRCGPGLGSCCESEEDYRP 480
Db 447 FKDSLIVQVTFDCACQACQQAEPFSPHRCCNGNGFECGVRCGPGLGSCCESEEDYRP 506
QY 481 SQDQESPRGQPVCSORGSCLCGCCVHSDFKGKIGDVSFSSTEAKVQSCPUERKEKFTIKVG 540
Db 507 SQDQESPRGQPVCSORGSCLCGCCVHSDFKGKIGDVSFSSTEAKVQSCPUERKEKFTIKVG 566
QY 541 QCSCGDCLCSDWTCYCNCTRTDTCMSSNGLCSGRGCECGSCVC1QPGSYGDTCEK 600
Db 567 QCSCGDCLCSDWTCYCNCTRTDTCMSSNGLCSGRGCECGSCVC1QPGSYGDTCEK 626
QY 601 CPTCPDACTKKECKKKDGRALHDEWCNRYCRDELEFSVKLDRGKDAVNTYNG 660
Db 627 CPTCPDACTKKECKKKDREPYMTENCRYCRDELESVKELDKDQKDAVNTYNG 686
QY 661 DDCVVRQFYEDSSGSKSLIVVVEEPCPKGD 692
Db 687 DDCVVRQFYEDSSGSKSLIVVVEEPCPKGD 718

PRIORITY: 1995-01-01

APPLICATION NUMBER: 08/218878

FILED: 13-JAN-1992

ATTORNEY/AGENT INFORMATION:

NAME: Lee, Wendy M.

REGISTRATION NUMBER: 00,000

REFERENCE/DOCKET NUMBER: P0552P1C3D1

TELECOMMUNICATION INFORMATION:

TELEPHONE: 415/225-1994

TELEFAX: 415/955-9881

TELEX: 910-371-7168

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 718 amino acids

TYPE: Amino Acid

TOPOLOGY: Linear

US-08-44490

FILED: 01-DEC-1989

PRIORITY: 1988-07-07

APPLICATION NUMBER: 07/290224

ATTORNEY/AGENT INFORMATION:

NAME: Lee, Wendy M.

REGISTRATION NUMBER: 00,000

REFERENCE/DOCKET NUMBER: P0552P1C3D1

TELECOMMUNICATION INFORMATION:

TELEPHONE: 415/225-1994

TELEFAX: 415/955-9881

TELEX: 910-371-7168

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 718 amino acids

TYPE: Amino Acid

TOPOLOGY: Linear

US-08-445-042-4

Query Match 90.7%; Score 3768; DB 1; Length 718; Best Local Similarity 99.3%; Pred. No. 2.5e-295; Matches 687; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 GPNICUTTRGVSSCQCLAVSPMCWCSDAALPLGSPRDLKLNKDNCAPESEIEPVSE 60

Db 27 GPNICUTTRGVSSCQCLAVSPMCWCSDAALPLGSPRDLKLNKDNCAPESEIEPVSE 86

QY 61 ARVLEDPLSLDKSGDSSQVTOQSPQRALRLRPLPDSKNSIQLQVROVEDYPVDIYIMD 120

Db 87 ARVLEDPLSLDKSGDSSQVTOQSPQRALRLRPLPDSKNSIQLQVROVEDYPVDIYIMD 146

QY 121 SYSKDDIWSIONGKTLATQMRKLTSLRIGFAVDKPVSQPMYISPPALELENPCYDM 180

Db 147 SYSKDDIWSIONGKTLATQMRKLTSLRIGFAVDKPVSQPMYISPPALELENPCYDM 206

QY 181 KTCCLPMPFGYKHYLTLDQTVTRENEEVYKOSYRNRAPEKGFDATMQATVYDEKIGWNR 240

Db 207 KTCCLPMPFGYKHYLTLDQTVTRENEEVYKOSYRNRAPEKGFDATMQATVYDEKIGWNR 266

QY 241 DASHLVLVFTDATHLADGLAGIYQPNQDQCHVGDDNHSASTIMDYPISGLMTEKLS 300

Db 267 DASHLVLVFTDATHLADGLAGIYQPNQDQCHVGDDNHSASTIMDYPISGLMTEKLS 326

QY 301 QKNINLFLAVTENVNLYQNYSLIPGTVGLSMDSNNVQLIVDAYGKRSKVELEVR 360

Db 327 OKNINLFLAVTENVNLYQNYSLIPGTVGLSMDSNNVQLIVDAYGKRSKVELEVR 386

QY 361 DLPEELSLSFNATCLNNREVIGPLKSCMGLKIGDTVSIEAVRGCPQEIKRSFTIRPVG 420

Db 387 DLPEELSLSFNATCLNNREVIGPLKSCMGLKIGDTVSIEAVRGCPQEIKRSFTIRPVG 446

QY 421 FKSLIYQVTFDCACAOQAQPNSHRCNNNGTPEGCVQCGPGWLGSGQCECSEEDYRP 480

Db 447 FKSLIYQVTFDCACAOQAQPNSHRCNNNGTPEGCVQCGPGWLGSGQCECSEEDYRP 506

QY 481 SQODECSPREGPVCOSORGECICLGQCVCHSSDFGKIGTKYECDDFSCVRKGEMCGSHG 540

Db 507 SQODECSPREGPVCOSORGECICLGQCVCHSSDFGKIGTKYECDDFSCVRKGEMCGSHG 566

RESULT 5

US-08-054-077C-2

Sequence 2, Application US/08054077C

Patent No. 552679

GENERAL INFORMATION:

APPLICANT: RAMASWAMY, HEMAVATHI

TITLE OF INVENTION: HUMAN INTEGRIN 5 SUBUNIT PROTEIN

NUMBER OF SEQUENCES: 7

CORRESPONDENCE ADDRESS:

ADDRESSEE: DAVID G. CONLIN; DIKE, BRONSTEIN, ROBERTS & CUSHMAN

STREET: 130 WATER STREET

CITY: BOSTON

STATE: MASSACHUSETTS

ZIP: 02109

COUNTRY: US

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-POS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/054,077C

FILED: 27 APR 1993

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/694314

FILED DATE: 01-MAY-1991

ATTORNEY/AGENT INFORMATION:

NAME: RESNICK, DAVID S.

REGISTRATION NUMBER: 34235

REFERENCE/DOCKET NUMBER: 40937

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617) 523-3400

TELEFAX: (617) 523-6440

TELEX: 200291 STRE UU

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 799 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-054-077C-2

Query Match 56.3%; Score 2339; DB 1; Length 799; Best Local Similarity 55.5%; Pred. No. 3.9e-180; Matches 433; Conservative 118; Mismatches 205; Indels 24; Gaps 11;

QY 1 GPNICUTTRGVSSCQCLAVSPMCWCSDAALPLGSP---RDLKLNKDNCAPESE 55

Db 24 GLNICKTSGSATCSCERCLLHPKACWCKE---DFFGSPRITSRSDDRLANLVNGCGGE 114

QY 56 FPVSEARVLEDPLSLDKSGDSS-QVTQSPQRALRLRPLPDSKNSIQLQVROVEDYPVDI 114

Db 81 SPASSFHLVSLPLSSKGSGSAGMDVQMTPOEAVNLRLPGDKITFQVQVROVEDYPVDI 140

QY 115 YLMDLSYSKDDIWSIONGKTLATQMRKLTSLRIGFAVDKPVSQPMYISPPALE 174

Patent No. 6339148
 GENERAL INFORMATION:
 APPLICANT: Sheppard, Dean
 APPLICANT: Quaranta, Vito
 APPLICANT: Pytelia, Robert
 TITLE OF INVENTION: A No. 6339148el Integrin Beta Subunit and Uses
 TITLE OF INVENTION: Thereof
 NUMBER OF SEQUENCES: 62
 CORRESPONDENCE ADDRESS:
 ADDRESSE: Townsend and Townsend and Crew LLP
 STREET: Two Embarcadero Center, Eighth Floor
 CITY: San Francisco
 STATE: California
 COUNTRY: USA
 ZIP: 94111-3834
 COMPUTER READABLE FORM:
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/938, 085A
 FILING DATE: 26-SEP-1997
 CLASSIFICATION: 435
 PRIORITY APPLICATION DATA:
 APPLICATION NUMBER: US 07/728, 215
 FILING DATE: 11-JUL-1991
 ATTORNEY/AGENT INFORMATION:
 NAME: Parent, Annette S.
 REGISTRATION NUMBER: 42, 058
 REFERENCE/DOCKET NUMBER: 023070-080210US
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (415) 576-0300
 TELEFAX: (415) 576-0300
 INFORMATION FOR SEQ ID NO: 27:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 788 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-938-085A-27

Query Match
 Best Local Similarity 51.9%; Score 2156; DB 4; Length 788;
 Matches 389; Conservative 133; Mismatches 231; Indels 16; Gaps 8;

QY 1 GPNICITR---GVSSQQCLAVSPMCAWCDEAL--PLG-SPRCDLKENILKDCAPE 52
 Db 14 GRNDSRTRWLCLGAGETCCLLPGOCWAQAOENFTPHPSGVGERCDTPANILAKGCQLN 73
 QY 53 SIEPPVSEARVLEDRPUSDKGSDSSQVTOVSQRAFLRRLRDDSKNFSIQVROVEDPYV 112
 Db 74 FIEPNPVSQVEILKNKPLSVGRQKNNSDVIQVAPQSLILKURGGAQTLQVAVROTEDPYV 133
 QY 113 DTYYLMDLSYSMKDDLWSIQNGLGKLTNSLRIGTGFVYDKPVSPSYMLSPPEA 172
 Db 134 DLYYLMDLSAAMDNLNTIKEGLGSLKEMSKLTSNFRLGFGSVERPKVSPFVW-TPPEE 192
 QY 173 LEPYCDYDMKTCLPMPFGYKHWLTQDQYTRFEVKKQSVRNDAPEGGDAIMAVC 232
 QY 193 IANPCCSITPYCLPTEGPKHLPLNDAERFNEIVKNOKIANIDAPPEGGDAIMQAVC 252
 Db 233 DEKIGWRNDASHILVFTIDAKTHIALDGRAGTQVQPDGQHVGSDONHYSATTMDYPSL 292
 Db 253 KERIGWRNDASHILVFWSDASHFGMDSKLAGIVPMDGLCHLDSKNEYSMTYLEPTI 312
 QY 293 GLMTEKLSOKNINLIFAVTNTVLYNQSELIPGTVWGVSMDSNVLOLIVDAYKIR 352
 Db 313 GOIJDKLKVQNVNLIFAVTQEQVHLYENYKLIPGATVGLLQDKSGNLTQLOLTISAYEELR 372
 QY 353 SKVELEYRDLPEELSFSNATCINNEVPIGLKSCMGKIGITVFSSTEAKRGCPEKEK 412

Query Match
 Best Local Similarity 44.1%; Score 1831; DB 2; Length 798;
 Matches 350; Conservative 134; Mismatches 268; Indels 28; Gaps 14;

RESULT 8
 US-07-728-215-30
 Sequence 30, Application US/07728215
 Patent No. 5962643
 GENERAL INFORMATION:
 APPLICANT: Sheppard, Dean
 APPLICANT: Quaranta, Vito
 APPLICANT: Pytelia, Robert
 TITLE OF INVENTION: A No. 5962643el Integrin Beta Subunit and Uses
 TITLE OF INVENTION: Thereof
 NUMBER OF SEQUENCES: 43
 CORRESPONDENCE ADDRESS:
 ADDRESSE: Pretty, Schroeder, Brueggemann & Clark
 STREET: 4370 La Jolla Village Drive, Suite 700
 CITY: San Diego
 STATE: California
 COUNTRY: United States of America
 ZIP: 92122
 COMPUTER READABLE FORM:
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/07/728, 215
 FILING DATE: 19910711
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Campbell, Kathryn A.
 REGISTRATION NUMBER: 31,815
 REFERENCE/DOCKET NUMBER: P31 8717
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (619) 535-9001
 INFORMATION FOR SEQ ID NO: 30:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 798 amino acids
 TYPE: AMINO ACID
 STRANDEDNESS: single
 TOPOLOGY: linear
 US-07-728-215-30

Db 620 CTDPKFQGOTCEMCQTCGLGCAEHKECVOCAERENKEKKD-TCTQECSYFNITKVESRD 677
 Qy 644 EL-KOTGKAVW-NOTYKNEEDCIVRFQYEDSGSKSYLIVVEREPECCKGPDPILVLLSV 701
 Db 678 KLPQFVQDPDVSHSKEKDWDYFTSVNGNEVMHVNEMVPECPGDPDIPIVAGVV 737
 Qy 702 GAILLIGLAALLIWKLITLHDKEFAKEERAKMDTANNPYKEATSTFTNTIYRG 761
 Db 738 AGIVLIGLALLIWKLMLTIDRFAKETKEKMNADTGENPIYSAVTTVNPYEG 797
 RESULT 10
 US-07-728-215-29
 ; Sequence 29, Application US/07728215
 ; Patent No. 5962643
 ; GENERAL INFORMATION:
 ; APPLICANT: Sheppard, Dean
 ; APPLICANT: Quaranta, Vito
 ; APPLICANT: Pytela, Robert
 ; TITLE OF INVENTION: A No. 5962643el Integrin Beta Subunit and Uses
 ; NUMBER OF SEQUENCES: 43
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Pretty, Schroeder, Bruggemann & Clark
 ; STREET: 4370 La Jolla Village Drive, Suite 700
 ; CITY: San Diego
 ; STATE: California
 ; COUNTRY: United States of America
 ; ZIP: 92122
 ; COMPUTER READABLE FORM:
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/07/728,215
 ; FILING DATE: 19910711
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Campbell, Kathryn A.
 ; REGISTRATION NUMBER: 311,815
 ; REFERENCE/DOCKET NUMBER: P31 8717
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (619) 535-9001
 ; TELEFAX: (619) 535-8949
 ; INFORMATION FOR SEQ ID NO: 29:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 577 amino acids
 ; TYPE: AMINO ACID
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-07-728-215-29

Query Match 40.3%: Score 1676; DB 2; Length 577;
 Best Local Similarity 51.1%; Pred. No. 6.5e-127; Matches 297; Conservative 102; Mismatches 174; Indels 8; Gaps 5;

Qy 121 SYSMKDDIWSIONLGKMLATQMRKLTNSLRGFGAIVDKPKSPYMWISPEALENCYDM 180
 Db 1 SASMDDDLINTKEGLSLMSKLNFRGFGSEVEKPSPEMK-TTPEEATANPSSI 59
 Qy 181 KTCCLPMPGKAVLTLTDQVTFNEEVKOSVNRDAPEGGFDATMQATVCDKEKGWRN 240
 Db 60 PYTICLPPFGKFLPLTNDAREFNEVVKQKISANIDNPEGFDAIMQAACKERIGWRN 119
 Qy 241 DASHLUVTTDAKTHTALDGRLAGIVQPNDOCHYGSNDHNSASTMDYPSLGMTEKLS 300
 Db 120 DSHLHLVVFVSDADSHFGMDSKLAGIVIPNDGLCHLDSKNEYSMSTVMEYPTIGQOLIDKV 179
 Qy 301 QNNNLFIAVENVVYQNSELICITVQVLSMDSSNVQQLIVDAYGIRSKVLEVR 360

RESULT 11
 US-08-938-085A-29
 ; Sequence 29, Application US/08938085A
 ; Patent No. 6339148
 ; GENERAL INFORMATION:
 ; APPLICANT: Sheppard, Dean
 ; APPLICANT: Quaranta, Vito
 ; APPLICANT: Pytela, Robert
 ; TITLE OF INVENTION: A No. 6339148el Integrin Beta Subunit and Uses
 ; NUMBER OF SEQUENCES: 62
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Townsend and Townsend and Crew LLP
 ; STREET: Two Embarcadero Center, Eighth Floor
 ; CITY: San Francisco
 ; STATE: California
 ; COUNTRY: USA
 ; ZIP: 94111-3834
 ; COMPUTER READABLE FORM:
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/938,085A
 ; FILING DATE: 26-SEP-1997
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/728, 215
 ; FILING DATE: 11-JUL-1991
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Parent, Annette S.
 ; REGISTRATION NUMBER: 42,058
 ; REFERENCE/DOCKET NUMBER: 023070-080210US
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (415) 576-0200
 ; INFORMATION FOR SEQ ID NO: 29:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 577 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-938-085A-29

Db 180 QNNVLLIFAVENVVYQNSELICITVQVLSMDSSNVQQLIVDAYGIRSKVLEVR 239
 Qy 361 DEPEELSLSENATCLNNENVIGLKSMSGKIGDVTFSLEAKVRGCPQCEKESTIKPGV 420
 Db 240 GDTENGLNLISLFSAVCNNGTLFPHQKCLHMKVGETASFNVNTVSIPNC-VERKSRRHVIKPGV 298
 Qy 421 FRSLSIVQVTFDCACQQAEPNSHRCNNONGTEFCGVRCGFWLGSOCCESEEDYRP 480
 Db 299 LGDTLELIVSBCSCBCKEVEVNNSKCHNGNSYQCVGACNPFGMCHCEGEGDTL-- 356
 Qy 481 SQDDESSPREQGPVCSSQRGECLCGQCVCHSSDFGKITGKVCCECDDFSVYKGMCSHG 540
 Db 357 -STDSEKETPDHPSCRGDGCGQCHIPLSPYGIYHGPYCQCNFSCTRHKGLGONG 415
 Qy 541 QCSGCGCICLSDWTVGWCNCNTTRPDMSSNGLCQSGRKCECSCVCIQPGSGDTEK 600
 Db 416 DCECGECVCRSRGWTGECNCNTSTDTCISEDGTLCSRGDCVCGKVCACNPGASGPTCER 475
 Qy 601 CPTCPDACTFKKECKKFDRGALHIDENTCNRCDETESVKELKD-GKD-AVNCTWKN 659
 Db 476 CPTCSBPCNSRSC-TECHLSDAQGQGE--CVDKCKLAGVTISREADSKDSSVSCS1QG 533
 Qy 660 EDDCVVRFQYEDSSGKILYVVEREPECPCPKGPDPILVLLSV 700
 Db 534 ENECLITFLISTDNECKTITHNISEKDCPKPPNPIMMLGV 574

RESULT 13
US-08-938-085A-33
SEQUENCE 33, Application US/08938085A
; Patent No. 6339148
GENERAL INFORMATION:
; APPLICANT: Shepard, Dean
; APPLICANT: Quaranta, Vito
; APPLICANT: Pytel, Robert
TITLE OF INVENTION: A novel Integrin Beta Subunit and Uses
TITLE OF INVENTION: Thereof
NUMBER OF SEQUENCES: 62
CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
ZIP: 94111-3334
COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-POS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/938,085A
; FILING DATE: 26-SEP-1997
; CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 07/728,215
; FILING DATE: 11-JUL-1991
ATTORNEY/AGENT INFORMATION:
; NAME: Parent, Annette S.
; REGISTRATION NUMBER: 412,058
; REFERENCE/DOCKET NUMBER: 023070-080210US
TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 33:
SEQUENCE CHARACTERISTICS:
; LENGTH: 846 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-938-085A-33

Query Match 37.0%; Score 1535.5; -DB 4; Length 846;
Best Local Similarity 38.7%; Pred. No. 2.3e-115;
Matches 320; Conservative 128; Mismatches 292; Indels 87; Gaps 21;

QY 4 ICTTRKGVSSQQCLAVSPMCAWCSEALPIGSPRDLKENLKNCAPESSIEFPVSEARV 63
; | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
; VSTCASEKCKHTCLOTEG-CAWCMOPDFK-GQSRQYQNTSSL--CPEEFAYSPBITVQI 91
; | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
; DB 64 LEDRPL-----SDKGSGDSS----- 78
; | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
; DB 92 LYNKKLTINQYKAELAAGGGGAMGSSSSSSSSSSSSYFQSSGSSSSASGEEYSAG 151
; | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
; QY 79 QVTOVSPORALRRLRPPDSKNSFSTQYRVEDPYPDIVYMLDSYMKDPLWSIONLGTL 138
; : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
; DB 152 EIVQIQPQSKRLALRNVNEKHNIKISYSQAEQYGPVDLYMLDLSKSMEDDKAKLSTLGDKL 211
; | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
; QY 139 ATQMKLTSURLRIGGAFVFKPVSPYMISSPPEALENPYDQYKVLITD 198
; | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
; DB 212 SETMRKTTNNPHLGFSGFVQKMLPYV-STIPKLEHC---ENCKAPYQYQNHMLNN 266
; | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
; QY 199 QVTRNEENKKQSVSRNRDAPEGGFDAMQATWCDKLGWRNDASHLUVFTDAKTHAL 258
; | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
; DB 267 NTESNEENKKATNSGNLDPAPEGGFDAIMQAIAGRSQQLWRQARQLLWFTDAGFHAG 326

RESULT 14
US-08-789-078-1
SEQUENCE 1, Application US/08789078
; Patent No. 5843885
GENERAL INFORMATION:
; APPLICANT: Benedict, Stephen H.
; APPLICANT: Slaahan, Texuna
; APPLICANT: Chan, Marcia
; APPLICANT: Tibbetts, Scott
TITLE OF INVENTION: ICAM-1/LFA-1 PEPTIDES FOR INDUCING
TITLE OF INVENTION: IMMUNE TOLERANCE
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
; ADDRESSEE: HOYER, WILLIAMS, TIMMONS & COLLINS
; STREET: 1101 Walnut St.
; CITY: Kansas City
; STATE: MO
; COUNTRY: USA
ZIP: 64106
COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/789,078
; FILING DATE: 03-FEB-1997
; CLASSIFICATION: 514
PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: 08/229, 513
; FILING DATE: 19-APR-1994
ATTORNEY/AGENT INFORMATION:
; NAME: Collins, John M.
; REGISTRATION NUMBER: 26262
; REFERENCE/DOCKET NUMBER: 22833

TELECOMMUNICATION INFORMATION: QY

TELEPHONE: (816)474-9050

TELEFAX: 816-474-9057

TELEX: 434-363

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS: QY

LENGTH: 769 amino acids

TYPE: amino acid

MOLECULE TYPE: protein

HYPOTHETICAL: NO

ORIGINAL SOURCE: QY

ORGANISM: Homo sapiens

TISSUE TYPE: Tonsil

FEATURE: QY

NAME/KEY: Region

LOCATION: 1..22

OTHER INFORMATION: /label= "signal sequence"

FEATURE: QY

NAME/KEY: Region

LOCATION: 449..496

OTHER INFORMATION: /label= "repeat"

OTHER INFORMATION: /note= "cysteine rich repeat"

FEATURE: QY

NAME/KEY: Region

LOCATION: 497..540

OTHER INFORMATION: /label= "repeat"

OTHER INFORMATION: /note= "cysteine rich repeat"

FEATURE: QY

NAME/KEY: Region

LOCATION: 541..581

OTHER INFORMATION: /label= "repeat"

OTHER INFORMATION: /note= "cysteine rich repeat"

FEATURE: QY

NAME/KEY: Region

LOCATION: 582..617

OTHER INFORMATION: /label= "repeat"

OTHER INFORMATION: /note= "cysteine rich repeat"

FEATURE: QY

NAME/KEY: Domain

LOCATION: 701..723

OTHER INFORMATION: /label= "trans

OTHER INFORMATION: /note= "transmembrane domain"

FEATURE: QY

NAME/KEY: Domain

LOCATION: 724..769

OTHER INFORMATION: /label= "cyto

OTHER INFORMATION: /note= "cytoplasmic domain"

PUBLICATION INFORMATION: QY

AUTHORS: Pigott, C. J.

AUTHORS: Power, D. J.

TITLE: LFA-1 Amino acid sequence (B2) (from human tonsil)

JOURNAL: The Adhesion Molecule Facts Book

PAGES: 96-96

DATE: 1993

RELEVANT RESIDUES IN SEQ ID NO: 1: FROM 1 TO 769

RELEVANT RESIDUES IN SEQ ID NO: 1: FROM 770 TO 798-1

RELEVANT RESIDUES IN SEQ ID NO: 1: FROM 799 TO 833

RELEVANT RESIDUES IN SEQ ID NO: 1: FROM 834 TO 886

RESULT 15

US-08-752-633-1

; Sequence 1, Application US/08752633

; Patent No. 5863689

; GENERAL INFORMATION:

; APPLICANT: Benedict, Stephen H.

; APPLICANT: Siahaan, Teruna

; APPLICANT: Chan, Marcia

; APPLICANT: Tibbets, Scott

; TITLE OF INVENTION: ICAM-1/LFA-1 PEPTIDES FOR INDUCING IMMUNE TOLERANCE

; NUMBER OF SEQUENCES: 19

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: HOVEY, WILLIAMS, TIMMONS & COLLINS

; STREET: 1101 Walnut St.

; CITY: Kansas City

; STATE: MO

; COUNTRY: USA

; ZIP: 64106

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC Compatible

OPERATING SYSTEM: PC-DOS/M-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/752,633

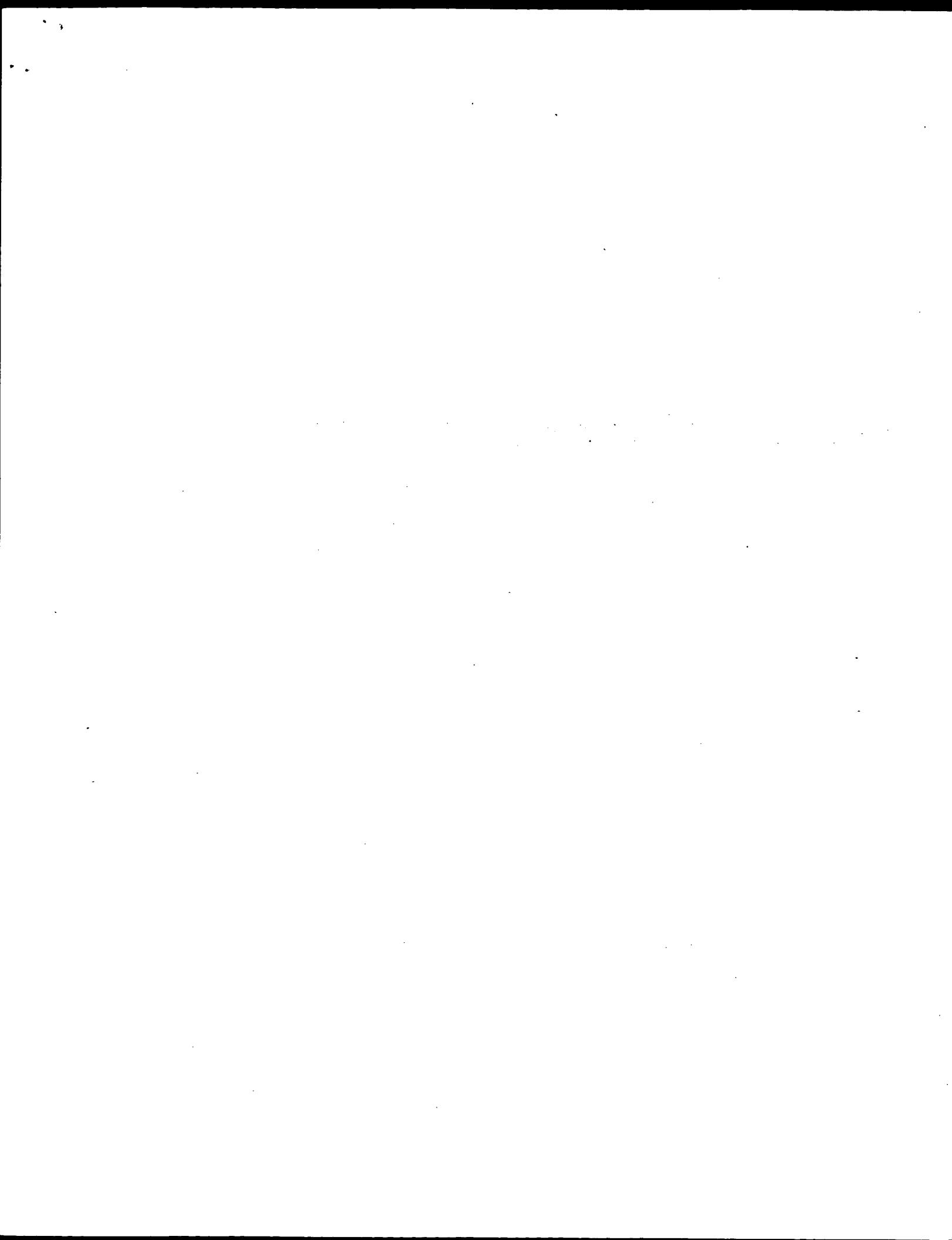
FILED DATE:

CLASSIFICATION: 514

ATTORNEY/AGENT INFORMATION: NAME: Collins, John M.
REGISTRATION NUMBER: 24262
REFERENCE/DOCKET NUMBER: 22833
TELECOMMUNICATION INFORMATION: TELEPHONE: (816) 474-9050
TELEFAX: 816 474-9057
TELEX: 434-363
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 769 amino acids
TYPE: amino acid
TOPOLOGY: both
NAME/KEY: Region
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
TISSUE TYPE: Tonsil
FEATURE: Region
LOCATION: 1..22
OTHER INFORMATION: /label= signal sequence
FEATURE: OTHER INFORMATION:
NAME/KEY: Region
LOCATION: 439..496
OTHER INFORMATION: /label= repeat
FEATURE: OTHER INFORMATION:
NAME/KEY: Region
LOCATION: 497..540
OTHER INFORMATION: /label= repeat
FEATURE: OTHER INFORMATION:
NAME/KEY: Region
LOCATION: 541..581
OTHER INFORMATION: /label= repeat
FEATURE: OTHER INFORMATION:
NAME/KEY: Region
LOCATION: 552..617
OTHER INFORMATION: /label= repeat
FEATURE: OTHER INFORMATION:
NAME/KEY: Domain
LOCATION: 701..723
OTHER INFORMATION: /label= trans
FEATURE: OTHER INFORMATION:
NAME/KEY: Domain
LOCATION: 724..769
OTHER INFORMATION: /label= cyto
OTHER INFORMATION: /note= "cytoplasmic domain"
PUBLICATION INFORMATION:
AUTHORS: Piquett,
AUGUST: Power,
TITLE: LFA-1 Amino acid sequence (B2) (from human
JOURNAL: The Adhesion Molecule Facts Book
PAGES: 96-96
DATE: 1993
RELEVANT RESIDUES IN SEQ ID NO: 1: FROM 1 TO 769

59 SEARVILEDPLSDKGSSQDSSQVTOVSPIRALRPPDSSKNSFSTQVRCVEDYDPIVYIM 118
Db 83 SLAEQEDHNGQK-----QISPKQVLYLRPGQAAFNVTFRPKGYPDLYIM 133
Db 119 DLSSYNSKKDLSWISQNGTQKLTQMKLTSNURGFGAFYKDPVSPYMSIPPEALENPY 178
Qy 134 DLSSYNSMDDLRNKKLGGDLRALNEITESSGRGFSYDVTWV-THPDQLRNPQ 192
Db 179 DMKTCILPQMGYKHLTLTQDVTRENEKQYSRNRRAPEGFDAMQATVKDQK 238
Qy 193 NKEKBCQPFPAFRVHLKLNTNSNPFQTEVGKOLISNLDAPEGFLDAMQVAACPELGW 252
Db 239 RNDASHLVLFTDAKHTIALDGRLAGIVQNDGQCHVGSDNHYASTTMDSPLGMLT 298
Qy 253 RN-VTRLVLFAFDGFFHFDGKLGALIFNDGCHL EDNLYKRSNFEDPSVQGLAH 310
Db 299 LSKOKINLNLIAFWTEWVNLYKQNSBLLPQGTYLIMPSNNVQOLIVAYGKRSKVE 358
Db 311 LAENNIQPIFAVTSMKVKYETLTIPIKPSAVGELSEDSSNVVHLKNAYNKLSSRFLD 370
Qy 359 VRDPEELSFNATCLNNEVIRPK-----SCMGKJGIVVSPSEAKVQCPDE 411
Db 371 HNALPDPTLKVYDPSQN----GVTHRNQPRGBCDQGVOINVITFQKVATECIG-- 423
Qy 412 KSFIFIKPVGKDSLWVQVFDGQACQAOAEPNSHRKCNNGNGTECCVGRCPGWLSSQC 471
Db 424 QSEFVIRALGFDPTWVWVQUPCQECRDS-LC-HGKGELECCGICRCGQYIEKNC 481
Db 472 ECSEBDYRPSQDQE--CSPREQGPVCSSORGECGCGQVCHSSDR-GK-ITGKCECDF 527
Db 482 ECQFQG--RSQSELEBCSCKRDNNSITC5G1GDCVGQCLHTSDVPGKJLJYQCECTIN 540
Qy 528 CVYRKGEMSSGHGQ--CSGDCLCLSDSMDWGYCNCCTRDTCTSSNGLCSERGKCCGS 585
Db 541 CERYNGQVCGGPGRGICFCGKCRCHPFGFEGSACOCERTTEGCLNPRVECGSGRGRCNV 600
Qy 586 CVCLQPGSTYGDTCBCKCPTCPDACTFKKECVECKFDRGALHDENTCNRYCRDEIESKEL 645
Db 601 CEC-HSGYOLPLCOPCPCGPGPKYIISCAECKJFEKPF--GKNCACPG----LQL 652
Db 646 KDTGKDAVNTYKNUEDCVYRFOY-YEDSSGGKSTLYVVEPEPCPKGPQDPLVWLSWGA 704
Db 653 SNNPVKGRICKERDSEGCWVAYLEQODMDRILYVDESRECVAAGPNIAAVGGTWAGI 712
Qy 705 LLIGLALLIWKLITHURKEFAFEERAKWDANNPLYKEAATSTFN 756
Db 713 VLIGILLVWIKALIHLSDLREYRFEKEKLKSQWN -DNPLFKSATWVMN 763

Search completed: May 19, 2002, 10:39:46
Job time: 9115 sec



GenCore - version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

On protein - protein search, using sw model

Run on: May 19, 2002, 09:42:13 ; Search time 55.89 Seconds
(without alignments)

Sequence: US-09-673-302a-1
1 GPNICTTRGVSSCQCLAVS.....NNPLYKEATSTFTNTIYRGP 762

Title: Perfect score: 4154
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 9609334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_71:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match Length	DB ID	Description
1	4154	100.0	788	2 I77349	platelet glycoprot
2	4146	99.8	778	2 A60798	platelet glycoprot
3	4119	99.2	788	2 A6547	platelet glycoprot
4	4038	97.2	753	2 B36268	platelet glycoprot
5	3577	86.1	723	2 PNU509	integrin beta-3 ch
6	3484	83.9	781	2 S43534	integrin beta-3 ch
7	3470	83.5	680	2 PNU510	integrin beta-3 ch
8	3260	78.5	708	2 A30308	integrin beta-3 su
9	3240	56.3	799	2 A7057	integrin beta-5 ch
10	2156	51.9	788	2 A7057	integrin beta-5 ch
11	1945	48.0	656	2 JC205	integrin, band 3 p
12	1945	48.5	803	1 I0CH3	integrin, band 3 p
13	1831	44.1	798	2 B2709	fibronectin receptor
14	1826	44.0	798	2 B2193	integrin beta-1*
15	1821	43.8	798	2 S01659	integrin beta-1 ch
16	1818	43.8	798	2 A82193	integrin beta-1 ch
17	1817	43.7	773	2 I46059	integrin beta-1, integrin su
18	1806	43.5	799	1 IJMSFB	fibronectin receptor
19	1801	43.4	799	2 JC126	integrin beta olig
20	1776	40.3	577	2 B37057	integrin beta-6 ch
21	1579	38.0	809	2 A57283	integrin beta ch
22	1553	37.4	772	2 S33659	integrin beta 2 ch
23	1547	37.3	846	2 A30889	integrin beta ch
24	1520	36.6	769	1 JC1121	leukocyte adhesion
25	1517	36.5	771	2 A45839	leukocyte adhesion
26	1515	36.5	770	2 S04847	leukocyte adhesion
27	1513	36.4	769	1 IJHOLM	leukocyte adhesion
28	1506	36.3	806	2 A46271	integrin beta-7 ch
29	1502	36.2	2 A40526	integrin beta-7 ch	

RESULT	1	ALIGNMENTS
177349	platelet glycoprotein IIIa precursor - human	
	N;Alternate names: GPIIIa	
	C;Species: Homo sapiens (man)	
	C;Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 04-Mar-2000	
	C;Accession: I77349; F56793	
	R;Frachet, P.; Uzan, G.; Thevenon, D.; Denarier, E.; Prandini, M.H.; Marguerie, G.	
	Nol. Biol. Rep. 14, 27-33, 1990	
	A;Title: GPIIb and GPIIIa amino acid sequences deduced from human megakaryocyte cDNAs	
	A;Reference number: 157461; PMID:9026533	
	A;Accession: I77349	
	A;Status: preliminary; translated from GB/EMBL/DDJB	
	A;Molecule type: mRNA	
	A;Residues: 1-788 <FRA>	
	A;Cross-references: GB:M35999; NID:9183532; PID:RAA35927; 1; PID:9306795	
	R;Catimel, B.; Parmentier, S.; Leung, L.L.; McGregor, J.L.	
	Biochem. J. 279, 419-425, 1991	
	A;Title: Separation of important new platelet glycoproteins (GPIA, GPIC, GPIC*, GPIIA	
	A;Reference number: A56793; MUID:9206194	
	A;Accession: F56793	
	A;Status: preliminary	
	A;Molecule type: protein	
	A;Residues: 27-30, 'X', 33-37 <CAT>	
	A;Experimental source: Platelet	
	C;Superfamily: integrin beta chain; laminin-type EGF-like homology	
	C;Keywords: glycoprotein; platelet	
	F;1-21/Domain: signal sequence #status predicted <SIG>	
	F;125,346,478,782/Binding site: carbohydrate (Asn) (covalent) #status predicted	
	Query Match	100.0%; Score 4154; DB 2; Length 788;
	Best Local Similarity	100.0%; Pred. No. 1.5e-256;
	Matches 762; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
	QY	1 GPNICTTRGVSSCQCLAVSPMCWCACSDPRLGSPRDLKENILKDNCAPESTIEFPVSE 60
	Db	27 GPNICTTRGVSSCQCLAVSPMCWCACSDPRLGSPRDLKENILKDNCAPESTIEFPVSE 86
	QY	61 ARVLEDPRPSDKGSDSSVTOVSPQRALLRDPDSDENSFIQVROQVEDYPVDIYVNDL 120
	Db	87 ARVLEDPRPSDKGSDSSVTOVSPQRALLRDPDSDENSFIQVROQVEDYPVDIYVNDL 146
	QY	121 SYSMKDDLWASIQNUGTKLATQMRKLTNSRIGGAFVQPKVPSYMSPPPEALENPYCDM 180
	Db	207 KTCCLPMFGKHWLTDQYTRFNEVKQKVSVNRDAPEGGFDIAQWVCDKIGRN 240
	QY	241 DASHLIVFTDAKTHIALDGLASIYQPNQGQCVIGSDNHYSASTMOPSLGMLTELS 300

Db	267	DASHLIVFTDAKTHIALDGRLAGIVQFNGOCHVGSNDHYASSTIMDYPISGLMTEKLS	326
Qy	301	OKNINLIFAVTENVNLYNQNSLEIPGTVGVLISMDSSNVNLQLIVDAYGKIRSKVLEVR	360
Db	327	OKNINLIFAVTENVNLYNQNSLEIPGTVGVLISMDSSNVNLQLIVDAYGKIRSKVLEVR	386
Qy	361	DLPEELSLISFNATCLNNEVITPLGKSCMQLIGDTVSIEAKVRGCQEKSKFTIKPGV	420
Db	387	DLPEELSLISFNATCLNNEVITPLGKSCMQLIGDTVSIEAKVRGCQEKSKFTIKPGV	446
Qy	421	FKDSLIVQVTFDCDCACACQAEQNSHRNNGNTFEGVCRCPGPGWIGSQECSEEDYR	480
Db	447	FKDSLIVQVTFDCDCACACQAEQNSHRNNGNTFEGVCRCPGPGWIGSQECSEEDYR	506
Qy	481	SQDDECSPREGQPVCSQREGCILGQCVHSSDFGKIKGKCYCDDDFSCVRYKGEMSGHG	540
Db	507	SQDDECSPREGQPVCSQREGCILGQCVHSSDFGKIKGKCYCDDDFSCVRYKGEMSGHG	566
Qy	541	QCSGDCDCSDMTGIVYCNCTRTDTOMSSNLICSGRKRCRGSCVCIQPSYGTCEK	600
Db	567	QCSGDCDCSDMTGIVYCNCTRTDTOMSSNLICSGRKRCRGSCVCIQPSYGTCEK	626
Qy	601	CPTCPDACTFKKCVCEKKFDRGALHIDNTCNYCRDIESEVKELDTGKAVNCTYNE	660
Db	627	CPTCPDACTFKKCVCEKKFDRGALHIDNTCNYCRDIESEVKELDTGKAVNCTYNE	686
Qy	661	DDCVVRFQYYEDSSGSKSLVYVVEPECPPKGPDILWLSVWGMGAILLIGLALLWKLIT	720
Db	687	DDCVVRFQYYEDSSGSKSLVYVVEPECPPKGPDILWLSVWGMGAILLIGLALLWKLIT	746
Qy	721	IHDKEFAKFEERARAKWDTANNPLVYKEATSTFTNITYRGT	762
Db	747	IHDKEFAKFEERARAKWDTANNPLVYKEATSTFTNITYRGT	788
RESULT	2		
A60798		platelet glycoprotein IIIa beta chain (version 2) - human (fragment)	
C		Alternative names: antigen CD61; integrin beta 3	
C		Species: Homo sapiens (man)	
C		Date: 30-Sep-1993 #sequence_revision: 30-Sep-1993 #text_change: 23-Jul-1999	
C		Accession: A60798; D3605; D32528; C32528; A36268	
R		Rosa, J.P.; Bray, P.F.; Gayet, O.; Johnston, G.I.; Cook, R.G.; Jackson, K.W.; Shuman, Blood 72, 593-600, 1988	
A		Title: Cloning of glycoprotein IIIa cDNA from human erythroleukemia cells and localization	
A		Reference number: A60798; MUID:8829434	
A		Status: not compared with conceptual translation	
A		Molecule type: mRNA	
A		Residues: 1-778 <Res>	
R		Lanza, F.; Kleffner, N.; Phillips, D.R.; Fitzgerald, L.A.	
J		Biol. Chem. 265, 18098-18103, 1990	
A		Title: Characterization of the human platelet glycoprotein IIIa gene. Comparison with	
A		Reference number: A36085; MUID:91009231	
A		Molecule type: DNA	
A		Accession: A36085	
A		Residues: 17-705, 'G', 707-778 <LAN>	
A		Cross references: GB:M5494	
R		Hiraiwa, A.; Matsusaga, A.; Shiku, H.; Takahashi, T.; Naito, K.; Yamada, K.	
Blood		69, 560-564, 1987	
A		Title: Purification and partial amino acid sequence of human platelet membrane glycoprotein IIIa	
A		Reference number: A90731; MUID:87101510	
A		Molecule type: protein	
A		Accession: 208-224 <H12>	
A		Accession: C32528	
A		Molecule type: protein	
A		Residues: 429-433 <H1R>	
R		Zimrin, A.B.; Gidwitz, S.; Lord, S.; Schwartz, E.; Bennett, J.S.; White II, G.C.; Ponc	
J		Biol. Chem. 265, 8590-8595, 1990	
A		Title: The genomic organization of platelet glycoprotein IIIa.	
A		Reference number: A36268; MUID:90256778	
RESULT	3		
A25547		platelet glycoprotein IIIa beta chain precursor (version 1) - human	

A: Accession: A36268
A: Molecule type: DNA
A: Residue: 18-'05, 'G', 707-778 <ZIM>
A: Cross references: GB:J04427
A: Note: the authors translated the codon GAT for residue 233 as Glu, GAT for residue 234 as Asp.
A: Gene: GDB:ITG3
A: Cross-references: GDB:120013; OMIM:173470
A: Map position: 17q21.32-17q21.32
C: Superfamily: integrin beta chain; laminin-type EGF-like homology
C: Keywords: alternative splicing; glycoprotein

Query Match 99.8%; Score 4146; DB 2; Length 778;
Best Local Similarity 99.8%; Pred. No. 4; 6e-256;
Matches 761; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Query 1 GPNICITRGVSSCOCIAVSPMCACSDAELPIGSPRCDIENLKDNCAPESLEPVE 60
Db 17 GPNICITRGVSSCOCIAVSPMCACSDAELPIGSPRCDIENLKDNCAPESIEFPVE 76

Query 61 ARVLEDPLSDKGSDSDSQTVQSPQRALRRLDDSKNSFIQVQVEDYPVIIYMLD 120
Db 77 ARVLEDPLSDKGSDSDSQTVQSPQRALRRLDDSKNSFIQVQVEDYPVIIYMLD 136

Query 121 SYSMKDDLWSTONLGTKLATOMRKTNSRIGAFVAFKPVSPWYIISPEALENPVCMD 180
Db 137 SYSMKDDLWSTONLGTKLATOMRKTNSRIGAFVAFKPVSPWYIISPEALENPVCMD 196

Query 181 KTTCLCPMFGKRVLTIDTQVTFNEEVKKQSVSNRDAPEGGDAIMQATVCDKEIGRN 240
Db 197 KTTCLCPMFGKRVLTIDTQVTFNEEVKKQSVSNRDAPEGGDAIMQATVCDKEIGRN 256

Query 241 DASHLIVFTDAKTHIALDGRLAGIVQFNGOCHVGSNDHYASSTIMDYPISGLMTEKLS 300
Db 257 DASHLIVFTDAKTHIALDGRLAGIVQFNGOCHVGSNDHYASSTIMDYPISGLMTEKLS 316

Query 301 OKNINLIFAVTENVNLYNQNSLEIPGTVGVLISMDSSNVNLQLIVDAYGKIRSKVLEVR 360
Db 317 OKNINLIFAVTENVNLYNQNSLEIPGTVGVLISMDSSNVNLQLIVDAYGKIRSKVLEVR 376

Query 361 DLPEELSLISFNATCLNNEVITPLGKSCMQLIGDTVSIEAKVRGCQEKSKFTIKPGV 420
Db 377 DLPEELSLISFNATCLNNEVITPLGKSCMQLIGDTVSIEAKVRGCQEKSKFTIKPGV 436

Query 421 FKDSLIVQVTFDCDCACACQAEQNSHRNNGNTFEGVCRCPGPGWIGSQECSEEDYR
 480 |

Db 437 FKDSLIVQVTFDCDCACACQAEQNSHRNNGNTFEGVCRCPGPGWIGSQECSEEDYR 496

Query 481 SQDDECSPREGQPVCSQREGCILGQCVHSSDFGKIKGKCYCDDDFSCVRYKGEMSGHG 540
Db 497 SQDDECSPREGQPVCSQREGCILGQCVHSSDFGKIKGKCYCDDDFSCVRYKGEMSGHG 556

Query 541 QCSGDCDCSDMTGIVYCNCTRTDTOMSSNLICSGRKRCRGSCVCIQPSYGTCEK 600
Db 557 QCSGDCDCSDMTGIVYCNCTRTDTOMSSNLICSGRKRCRGSCVCIQPSYGTCEK 616

Query 601 CPTCPDACTFKKCVCEKKFDRGALHIDNTCNYCRDIESEVKELDTGKAVNCTYNE 660

Db 617 CPTCPDACTFKKCVCEKKFDRGALHIDNTCNYCRDIESEVKELDTGKAVNCTYNE 676

Query 661 DDCVVFQYYEDSSGSKSLVYVVEPECPPKGPDILWLSVWGMGAILLIGLALLWKLIT 720
Db 677 DDCVVFQYYEDSSGSKSLVYVVEPECPPKGPDILWLSVWGMGAILLIGLALLWKLIT 736

Query 721 IHDKEFAKFEERARAKWDTANNPLVYKEATSTFTNITYRGT 762
Db 737 IHDKEFAKFEERARAKWDTANNPLVYKEATSTFTNITYRGT 778

N; Alternate names: endothelial cell glycoprotein IIIa
 C; Species: Homo sapiens (man)
 C; Date: 05-Oct-1988 #sequence_revision 05-Oct-1988 #text_change 20-Aug-1999
 C; Accession: A26547; A32332; S65439
 J. Biol. Chem. 262, 3936-3939, 1987
 A; Title: Protein sequence of endothelial glycoprotein IIIa derived from a cDNA clone. Ic
 A; Reference number: A26547; MUID:87165991
 A; Accession: A26547
 A; Molecule type: mRNA
 A; Residues: 1-708 <XMT>
 A; Cross-references: GB:J02703; NID:9183452; PIDN:AAA52589.1; PID:9306786
 R. Zimirin, A.B.; Elsman, R.; Vilaine, G.; Schwartz, E.; Bennett, J.S.; Ponce, M.
 J. Clin. Invest. 81, 1470-1475, 1988
 A; Title: Structure of platelet glycoprotein IIIa. A common subunit for two different membrane proteins
 A; Reference number: A32532; MUID:88213696
 A; Accession: A32532
 A; Molecule type: mRNA
 A; Residues: 1-111788 <ZIM>
 A; Cross-references: GB:MD0311; NID:9190107; PIDN:AAA60122.1; PIB:9190108
 R. Makagonenko, E.M.; Yakubenko, V.P.; Ingham, K.C.; Meaved, L.V.
 Eur. J. Biochem. 237, 205-211, 1996
 A; Title: Thermal stability of individual domains in platelet glycoprotein IIIb/IIIa.
 A; Reference number: S65437; MUID:96203926
 A; Accession: S65439
 A; Molecule type: protein
 A; Residues: 27-375414 <MAK>
 C; Superfamily: Integrin beta chain; laminin-type EGF-like homology
 C; Keywords: alternative splicing; cytoskeleton; glycoprotein; transmembrane protein
 F; 1-267(Domain: signal sequence #status predicted <SIG>
 F; 719-788/Product: platelet glycoprotein IIIa beta chain #status predicted <MAT>
 F; 719-747(Domain: transmembrane #status predicted <TM>
 F; 125, 346, 397, 478, 585, 680, 782/Binding site: carbohydrate (Asn) (covalent) #status predicted
 F; 125, 346, 397, 478, 585, 680, 782/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 99.2%; Score 4119; DB 2; Length 788;
 Best Local Similarity 99.3%; Pred. No. 2, 5e-254; Matches 757; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 GPNICTRGVSSCQCLAVSPMCAMCSDEALPLGSPRCDLKENLKDNCAPESTEPVSE 60
 Db 27 GPNICTRGVSSCQCLAVSPMCAMCSDEALPLGSPRCDLKENLKDNCAPESTEPVSE 60
 Qy 61 ARYLEDRPLSKGSGDSSQVQSPORIALRRLPDKSKNFSIQVRQVEDYFVDTYMLD 120
 Db 87 ARYLEDRPLSKGSGDSSQVQSPORIALRRLPDKSKNFSIQVRQVEDYFVDTYMLD 146
 Qy 121 SWSKDDLWSQNGLGKLTAKTOMRKLSNLRLGFGAFVDKPVSPYMTSPPEALEPCYDM 180
 Db 147 SWSKDDLWSQNGLGKLTAKTOMRKLSNLRLGFGAFVDKPVSPYMTSPPEALEPCYDM 206
 Qy 181 KTCIPLPMGKHYVHLTLTDQVTRFNEEKVQSVRNRDAPEGFDAIMQATVCDENIGWRN 240
 Db 207 KTCIPLPMGKHYVHLTLTDQVTRFNEEKVQSVRNRDAPEGFDAIMQATVCDENIGWRN 266
 Qy 241 DASHLVLVFTDAKTHIALDGRLAGIYQPNQGQHVGSDNHSASTMIDYPSLGMTEKLS 300
 Db 267 DASHLVLVFTDAKTHIALDGRLAGIYQPNQGQHVGSDNHSASTMIDYPSLGMTEKLS 326
 Qy 301 OKINLNLFAVTEWVNVLYQNKSLPRTGTVGSLMSSNVYQLLIVDAYGKTRSKYLEVR 360
 Db 327 QKINLNLFAVTEWVNVLYQNKSLPRTGTVGSLMSSNVYQLLIVDAYGKTRSKYLEVR 386
 Qy 361 DLPBELSFSNATCLNNNEVPIPGKSKMKGKIGDTVSFSTEAKVRGCPQEKEKSTIKPVG 420
 Db 387 DLPBELSFSNATCLNNNEVPIPGKSKMKGKIGDTVSFSTEAKVRGCPQEKEKSTIKPVG 446
 Qy 421 FKOSLIVQVTFDCDCACOQAEEBNSHRCCNNGCTFEGGVCRGGPQGNGSQCSEEDYRP 480
 Db 447 FKOSLIVQVTFDCDCACOQAEEBNSHRCCNNGCTFEGGVCRGGPQGNGSQCSEEDYRP 506
 Qy 481 SQQDECSPREGQVCSRGECIUGOCYCHSSPQGKTKGKTYCBBDDFSCVRYKGEMSGHG 540

RESULT 4

B36268 Platelet glycoprotein IIIa-II - human (fragment)
 N; Alternate names: integrin beta-3 chain
 C; Species: Homo sapiens (man)
 C; Date: 18-Jan-1991 #sequence_revision 13-Jan-1993 #text_change 20-Aug-1999
 C; Accession: B36268; A33907
 R. Zimirin, A.B.; Glidwitz, S.; Lord, S.; Schwartz, E.; Bennett, J.S.; White II, G.C.; P. J. Biol. Chem. 265, 8590-8595, 1990
 A; Title: The genomic organization of platelet glycoprotein IIIa.
 A; Reference number: A36268; MUID:90256778
 A; Accession: B36268
 A; Molecule type: protein
 A; Status: preliminary
 A; Residues: 1-753 <ZIM>
 A; Cross-references: GB:J05427
 A; Note: the authors translated the codon GAT for residue 216 as Glu, GAT for residue R. van Kuppevelt, T.H.M.S.M.; Languino, L.R.; Gallit, J.O.; Suzuki, S.; Ruoslahti, E. Proc. Natl. Acad. Sci. U.S.A. 86, 5415-5418, 1989
 A; Title: An alternative cytosolic domain of the integrin beta-3 subunit.
 A; Reference number: A33907; MUID:89315807
 A; Status: preliminary
 A; Molecule type: mRNA
 A; Residues: 680-753 <VAN>
 A; Cross-references: GB:W45108; NID:9186502; PIDN:AAA36121.1; PID:9386833
 C; Superfamily: integrin beta chain; laminin-type EGF-like homology
 C; Keywords: cell adhesion; glycoprotein

Query Match 97.2%; Score 4038; DB 2; Length 753;
 Best Local Similarity 99.6%; Pred. No. 3, 3e-249; Matches 740; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 PNICTRGVSSCQCLAVSPMCAMCSDEALPLGSPRCDLKENLKDNCAPESTEPVSEA 61
 Qy 2 PNICTRGVSSCQCLAVSPMCAMCSDEALPLGSPRCDLKENLKDNCAPESTEPVSEA 61
 Db 1 PNICTRGVSSCQCLAVSPMCAMCSDEALPLGSPRCDLKENLKDNCAPESTEPVSEA 60
 Qy 62 RYLEDRPLSKGSGDSSQVQSPORIALRRLPDKSKNFSIQVRQVEDYFVDTYMLD 121
 Db 61 RYLEDRPLSKGSGDSSQVQSPORIALRRLPDKSKNFSIQVRQVEDYFVDTYMLD 120
 Qy 122 YSMKDDLWSQNGLGKLTAKTOMRKLSNLRLGFGAFVDKPVSPYMTSPPEALEPCYDM 181
 Db 121 YSMKDDLWSQNGLGKLTAKTOMRKLSNLRLGFGAFVDKPVSPYMTSPPEALEPCYDM 180
 Db 61 RYLEDRPLSKGSGDSSQVQSPORIALRRLPDKSKNFSIQVRQVEDYFVDTYMLD 120
 Qy 182 TTCLPMGKHYVHLTLTDQVTRFNEEKVQSVRNRDAPEGFDAIMQATVCDENIGWRN 241
 Db 181 TTCLPMGKHYVHLTLTDQVTRFNEEKVQSVRNRDAPEGFDAIMQATVCDENIGWRN 240
 Qy 242 ASHLVLVFTDAKTHIALDGRLAGIYQPNQGQHVGSDNHSASTMIDYPSLGMTEKLS 301
 Db 241 ASHLVLVFTDAKTHIALDGRLAGIYQPNQGQHVGSDNHSASTMIDYPSLGMTEKLS 300

QY	302 KNINLIFAVTENVNLYQNYSELIPGTVGVLMSMDSSNVQLTVDAYGKIRSKVELYRD	361	Db	241 HYSASTLMDYPSLGLIMTEKLSQRNINLIFAVTENVNLYQNYSELIPGTVGVLSDSSN	300
Db	301 KNINLIFAVTENVNLYQNYSELIPGTVGVLMSMDSSNVQLTVDAYGKIRSKVELYRD	360	QY	340 VIQLIVDAYGKIRSKVELEVDLPEELSLSFNATCINEVTPGKSCMGKIGKGDVFSI	399
QY	362 LPEELSLSFNATCINEVTPGKSCMGKIGKGDVFSIPEELSLSFNATCINEVTPGKSCMGKIGKGDVFSI	421	Db	301 VIQLIVDAYGKIRSKVELEVDLPEELSLSFNATCINEVTPGKSCMGKIGKGDVFSI	360
Db	361 LPEELSLSFNATCINEVTPGKSCMGKIGKGDVFSIPEELSLSFNATCINEVTPGKSCMGKIGKGDVFSI	420	QY	400 EAKVRGGPQEKEKSFTIKPVFKDSLIVQVTFDCDCACQAEPMSHRCNGNTFEGV	459
QY	422 KDSLIVQVTFDCACQAEPMSHRCNGNTFEGVCRCPGWLGSQCESEEDYRPS	481	Db	361 EAKVRGGPQEKSFTIKPVFKDSLIVQVTFDCDCACQAEPMSHRCNGNTFEGV	420
Db	421 KDSLIVQVTFDCACQAEPMSHRCNGNTFEGVCRCPGWLGSQCESEEDYRPS	480	QY	460 CRCGPWGLGSQCECSEEDYRPSQDSCSPRREGQPVCSRQEGCLOGQCVCHSSDPIKGK	519
QY	482 QDQCSPRREGQPVCSQREBLCCQCVCHSSDPIKGK	541	Db	421 CRCQDWGLGSQCECSEEDYRPSQDSCSPRREGQPVCSRQEGCLOGQCVCHSSDPIKGK	480
Db	481 QDQCSPRREGQPVCSQREBLCCQCVCHSSDPIKGK	540	QY	520 YCECDPSFCVYKGEMSGHQCSQGCLCSDWVGYCNCCTRTDTCMSSNGLCSCRG	579
QY	542 CSGCGLCDSDWVGYCNCCTRTDTCMSSNGLCSCRGKCEGKCGSCVCIQPSYGDTEKC	601	Db	481 YCECDPSFCVYKGEMSGHQCSQGCLCSDWVGYCNCCTRTDTCMSSNGLCSCRG	540
Db	541 CSGCGLCDSDWVGYCNCCTRTDTCMSSNGLCSCRGKCEGKCGSCVCIQPSYGDTEKC	600	QY	580 KCECGSCVCIQPSYGDTCBECXCPDACTFKKECCKFDRGALHENTCNRYCDBI	639
QY	602 PTCBDACTFKKECVECKFDRGALHENTCNRYCDBE TESTVELKJRGKDADVNTYKED	661	Db	541 NCECGSCVCIQPSYGDTCBECXCPDACSFKKDCVECKFNRGSSYMMENNCRCDI	600
Db	601 PTCBDACTFKKECVECKFDRGALHENTCNRYCDBE TESTVELKJRGKDADVNTYKED	660	QY	640 ESKVKEKDGTGDAVNTYKEDDCVYRFOYEDSSGKSLVYVPEPCKGPDILWVLS	699
QY	662 DCVYRFOYEDSSGKSLVYVPEPCKGPDILWVLSUMGATLGLAALIWKLLITI	721	Db	601 EKVKEKDGTGDAVNTYKEDDCVYRFOYEDSSGKSLVYVPEPCKGPDILWVLS	660
Db	661 DCVYRFOYEDSSGKSLVYVPEPCKGPDILWVLSUMGATLGLAALIWKLLITI	720	QY	700 VMGATLGLAALIWKLLITIHKRKEFEEERAKRATNPNLYKEATSTFTNTY	759
QY	722 HKRKEFEEERAKRATNPNLYKEATSTFTNTY	744	Db	661 VMGATLGLAALIWKLLITIHKRKEFEEERAKRATNPNLYKEATSTFTNTY	720
Db	721 HKRKEFEEERAKRATNPNLYKEATSTFTNTY	743	QY	760 RGT 762	
Db	721 RGT 723		Db	721 RGT 723	
RESULT	5		RESULT	6	
PN0509	integрин beta-3 chain - rat (fragment)		PN0534	integрин beta3 - chicken	
C;Species: Rattus norvegicus (Norway rat)	C;Species: Gallus gallus (chicken)		C;Species: Gallus gallus (chicken)	C;Species: Gallus gallus (chicken)	
C;Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 23-Jul-1999	C;Date: 07-Sep-1994 #sequence_revision 26-May-1995 #text_change 20-Aug-1999		C;Date: 07-Sep-1994 #sequence_revision 26-May-1995 #text_change 20-Aug-1999	C;Date: 07-Sep-1994 #sequence_revision 26-May-1995 #text_change 20-Aug-1999	
C;Accession: PN0509	C;Accession: S33534; A49312		C;Accession: S33534; A49312	C;Accession: S33534; A49312	
B;Cleutat, A.M.; Rosa, J.P.; Letourneau, F.; Ponz, M.; Rifat, S.	R;Mirura, H.; Cao, X.; Ross, F.P.; Chiba, M.; Teitelbaum, S.L.		R;Mirura, H.; Cao, X.; Ross, F.P.; Chiba, M.; Teitelbaum, S.L.	R;Mirura, H.; Cao, X.; Ross, F.P.; Chiba, M.; Teitelbaum, S.L.	
A;Title: A comparative analysis of cDNA-derived sequences for rat and mouse beta-3 integrin	A;Title: 1,25-dihydroxyvitamin D(3) transcriptionally activates the beta(3)-integrin		A;Title: 1,25-dihydroxyvitamin D(3) transcriptionally activates the beta(3)-integrin	A;Title: 1,25-dihydroxyvitamin D(3) transcriptionally activates the beta(3)-integrin	
A;Reference number: PN0509; MUID:93290675	A;Reference number: S43534; MUID:94164000		A;Reference number: S43534; MUID:94164000	A;Reference number: S43534; MUID:94164000	
A;Accession: PN0509	A;Status: nucleic acid sequence not shown		A;Status: preliminary	A;Status: preliminary	
A;Molecule type: mRNA	A;Molecule type: mRNA		A;Molecule type: mRNA	A;Molecule type: mRNA	
A;Residues: 1-723 <CITE>	A;Residues: 1-723 <CITE>		A;Cross-references: EMBL:Y72378; PID:9474038; PID:CAA51069.1; PID:9474039	A;Cross-references: EMBL:Y72378; PID:9474038; PID:CAA51069.1; PID:9474039	
Query Match 86.1%; Score 3577; DB 2; Length 723;	Query Match 83.9%; Score 3484.5; DB 2; Length 781;		A;Note: the authors translated the codon TTT for residue 174 as GY, GGC for residue 174	A;Note: the authors translated the codon TTT for residue 174 as GY, GGC for residue 174	
Best Local Similarity 89.9%; Pred. No. 6.7e-220; Matches 650; Conservative 42; Mismatches 31; Indels 0; Gaps 0;	Best Local Similarity 83.0%; Pred. No. 5.6e-214; Matches 632; Conservative 58; Mismatches 70; Indels 1; Gaps 1;		R;Cao, X.; Ross, F.P.; Zhang, L.; Macdonald, P.N.; Chappel, J.; Teitelbaum, S.L.	R;Cao, X.; Ross, F.P.; Zhang, L.; Macdonald, P.N.; Chappel, J.; Teitelbaum, S.L.	
QY	40 LRENLLKDCAPESTIPEVPSARVLDRPLSDKGSDSSVTOVSPORTAHLRPPDSK 99		J; Biol. Chem. 268: 2737-27380, 1993	J; Biol. Chem. 268: 2737-27380, 1993	
Db	1 IKNKLKDCKCSPESIPEVPSARVLDRPLSDKGSDSSVTOVSPORTAHLRPPDSK 60		A;Title: Cloning of the promoter for the avian integrin beta-3 subunit gene and its regulation	A;Title: Cloning of the promoter for the avian integrin beta-3 subunit gene and its regulation	
QY	100 PFIQVROYEDPVDTYLMDSYMDLSDMDSIONGKLTATQMRKLTNSRIGCAFVDK 159		A;Reference number: A49312; MUID:94086557	A;Reference number: A49312; MUID:94086557	
Db	61 FSLQVQROVEDPVDTYLMDSYMDLSDMDSIONGKLTATQMRKLTNSRIGCAFVDK 120		A;Accession: A49312	A;Accession: A49312	
QY	160 PVPSPYMPISPAEALNPYDVKITCIPMFGKHKVLTLDQVNEEVKQSVNRDAP 219		A;Molecule type: DNA	A;Molecule type: DNA	
Db	121 PVPSPYMPISPAEALNPYDVKITCIPMFGKHKVLTLDQVNEEVKQSVNRDAP 180		A;Residues: 1-26 <CITE>	A;Residues: 1-26 <CITE>	
QY	220 EGGFDATMQATWCDEKIGWRNDASHLIVFTDAKTHIALDGRLAGIVOPNDGCHVSDN 279		A;Cross-references: GB:X75348; PID:9411308; PID:CAA53095.1; PID:9452854	A;Cross-references: GB:X75348; PID:9411308; PID:CAA53095.1; PID:9452854	
Db	181 EGGFDATMQATWCDEKIGWRNDASHLIVFTDAKTHIALDGRLAGIVOPNDGCHVSDN 240		C;Super-family: integrin beta chain; laminin-type EGF-like homology	C;Super-family: integrin beta chain; laminin-type EGF-like homology	
QY	280 HYSASTLMDYPSLGLIMTEKLSQRNINLIFAVTENVNLYQNYSELIPGTVGVLSDSSN 339		Query Match 83.9%; Score 3484.5; DB 2; Length 781;	Query Match 83.9%; Score 3484.5; DB 2; Length 781;	
Db	80 ITVLEPLSLDKGGGGST-TTOMSPORIQLRLRPPDSKNSFQVROVEDYPVDTYLM 138		Best Local Similarity 83.0%; Pred. No. 5.6e-214; Matches 632; Conservative 58; Mismatches 70; Indels 1; Gaps 1;	Best Local Similarity 83.0%; Pred. No. 5.6e-214; Matches 632; Conservative 58; Mismatches 70; Indels 1; Gaps 1;	

Oy 121 SYSMKDLWSIONGLTAKLATOMRKLTSNRLIGEAFVDKPVSPWMVISPEALENPXYDM 180
 Db 139 SNSMKDLKNTQNLGTLKASMRKLTSNRLIGEAFVDKPVSPWMVISPEAKINPCYBI 198
 Oy 181 KITCLPMFGKHKVLTQDQTRFNEVKOSVRNDAEFGFAIMQATVCDERKIGRN 240
 Db 199 GEKCLPMFGKHKVLTQDQTRFNEVKOSVRNDAEFGFAIMQATVCDERKIGRN 258
 Oy 241 DASHLIVFTTDAKTHALDGRLAGLQVOPNDQCHGSDNHYSASTMDYPSLGMTEKLS 300
 Db 259 DASHLIVFTTDAKTHALDGRLAGLQVOPNDQCHGSDNHYSASTMDYPSLGMTEKLS 318
 Oy 301 QKNTNLIFAVTENVNLYNQSELIPGTVGVLSDSSNVLQLTVDAYGKIRSKVELEV 360
 Db 319 QKNTNLIFAVTENVNLYNQSELIPGTVGVLSDSSNVLQLTVDAYGKIRSKVELEV 378
 Oy 361 DLPEELSLSFNATCINNEVPGLSCMGKLGIDPVSESTEAKVGPQKEKSTIKPG 420
 Db 379 DLPEELSLSFNATCINNEVPGLSCMGKLGIDPVSESTEAKVGPQKEKSTIKPG 438
 Oy 421 FKDSLIVQVTPDCACQAOAEPNSHRCNNGFVFCGVCRCGWLGSQCESEEDYRP 480
 Db 439 FKDSLIVVWVNDGNCSCDAEANSFCNSKNGSLECGVRCNGRLGHCEDSEEANP 498
 Oy 481 SQODECSPREQOPVQESORGCLCQGCVCHSSDFEKITGKVCCECDFSCVRYKGMC 540
 Db 499 SEQDNQSPQGPQQLCSQRCGICGCVCHGSDFGKTVGKCECDFSCVRFKGMC 558
 Oy 541 QCSGGCCLCSDWITGVCYCNCITRDTGMSNGLUCSGRGKCECCSCVCTQPGSGDTC 600
 Db 559 QCSGGCCLCSDWITGVCYCNCITRDTGMSNGLUCSGRGKCECCSCVCTQPGSGDTC 618
 Oy 601 CPTCPDACTKEKCVCKKEKDRGALIDENTCNRCRDETESVKELKDTGDAVNCTYKE 660
 Db 619 CPTCSDACTKEKCVCKKEKDRGALIDENTCNRCRDETESVKELKDTGDAVNCTYKE 678
 Oy 661 DDCCVRQYEDSGSKSILVYVERBCKPGDILVLLSYMGATLIGAALLIWKLIT 720
 Db 679 NDCVWRFQYEDSGSKSILVYTERFCKPGDILVLLSYMGATLIGAALLIWKLIT 738
 Oy 721 IHDKRKFAKEFEEARAKWDTANPPLYKEATSTFNITRYG 761
 Db 739 IHDKRKFAKEFEEARAKWDTANPPLYKEATSTFNITRYG 779

RESULT 7
 PN0510
 C;Species: Mus musculus (house mouse)
 C;Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 23-Jul-1999
 C;Accession: IN0510
 R;Cleutier, A.M.; Rosa, J.P.; Letourneur, F.; Ponzc, M.; Rifat, S.
 Bichem. Biophys. Res. Commun. 193, 771-778, 1993
 A;Title: A comparative analysis of cDNA derived sequences for rat and mouse beta-3 integrin
 A;Reference number: PN0510
 A;Accession: PN0509; MUID: 93290675
 A;Status: nucleic acid sequence not shown
 A;Molecule type: mRNA
 A;Residues: 1-680 <CITE>
 C;Superfamily: integrin beta chain; laminin-type EGF-like homology
 C;Keywords: cell adhesion; duplication; heterodimer; membrane protein

Db 61 ASMRKLTSNRLIGEAFVDKPVSPWMVISPEAKINPCYBNACLPMFGKHKVLTQD 120
 Oy 199 QVTRPNEEVKOSVRNDAEFGFAIMQATVCDERKIGRN 258
 Db 121 QVSRENEEVKOSVRNDAEFGFAIMQATVCDERKIGRN 258
 Oy 259 DGRLAGTVOPNDQCHGSDNHYSASTMDYPSLGMTEKLS 318
 Db 181 DGRLAGTVOPNDQCHGSDNHYSASTMDYPSLGMTEKLS 318
 Oy 319 QNSELIGTIVKLSDDSSNVLQLTVDAYGKIRSKVELEV 378
 Db 241 QNSELIGTIVKLSDDSSNVLQLTVDAYGKIRSKVELEV 378
 Oy 379 VIGLKSOMGKLGIDPVSESTEAKVGPQKEKSTIKPG 438
 Db 301 VIGLKSOMGKLGIDPVSESTEAKVGPQKEKSTIKPG 438
 Oy 439 AQAEPNSHRCNNGTFCGVCRCGWLGSQCESEEDYRP 480
 Db 361 AFAQPSSPRCNGNGTFCGVCRCGWLGSQCESEEDYRP 480
 Oy 499 GECIGQCVCHSSDFEKGKTCYCECDFSCVRYKGMC 558
 Db 421 GECLGQCVCHSSDFEKGKTCYCECDFSCVRYKGMC 558
 Oy 559 NCTTRDTCMSNSNLCSGRGKCECGSCVCTQPGSYGDTCKECPDCKECPDCKECPDCKECK 618
 Db 541 KENKGTLHENTCSRICDDEQKELTDGKNAVNCTYKNEEDDCVRFQYEDTSGRAV 600
 Oy 679 LYVVEEPCKPGDILVLLSYMGATLIGAALLIWKLIT 738
 Db 601 LYVVEEPCKPGDILVLLSYMGATLIGAALLIWKLIT 738
 Oy 739 WDTRNNPPLYKEATSTFNIT 758
 Db 661 WDTRNNPPLYKEATSTFNIT 680

RESULT 8
 15130
 Integrin beta-3 subunit - African clawed frog
 C;Species: Xenopus laevis (African clawed frog)
 C;Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 20-Aug-1999
 C;Accession: 15130
 R;Ransom, D.G.; Hens, M.D.; Desimone, D.W.
 Dev. Biol. 160, 265-275, 1993
 A;Title: Integrin expression in early amphibian embryos: cDNA cloning and characterization
 A;Reference number: 15130; MUID: 94040374
 A;Accession: 15130
 A;Molecule type: mRNA
 A;Status: preliminary; translated from GB/EMBL/DDBJ
 A;Residues: 1-75 <RAN>
 A;Cross-references: GB:113591; NID:9467811; PID:9467812
 C;Superfamily: integrin beta chain; laminin-type EGF-like homology

Query Match 78.5%; Score 3260.5%; DB 2; Length 788;
 Best Local Similarity 75.9%; Pred. No. 1e-213;
 Matches 578; Conservative 89; Mismatches 94; Indels 1; Gaps 1;
 Oy 1 GPNICTTROVSSOCQCLAVSPMCWCSDAELPLGSPRCDKLENKNCAPESTIEPVE 60
 Db 24 GASICATGQVSSCQCLSVPOCACWSQEVFGKGAPRCDLKSILLSNCELKIEFPVST 83
 Oy 61 ARVLEDPLSKGS-GDSSQVQYSPQRLRRLRDPDKNSFSTQVQVEDYFDYIYMD 119
 Db 84 VLVVENPLSVKSGSEGDRREITOMSPKIDLERPDKTFVNLQVROVEDYFDYIYMD 143

Query 79 QVTPQVSPORTAHLRDLDSKNSFSTQVQVEDYFDYIYMDLSYSMKDDWISQNLGTL 138
 Db 1 QVTPQVSPORTAHLRDLDSKNSFSTQVQVEDYFDYIYMDLSYSMKDDWISQNLGTL 60
 Oy 139 ATQMRKITSNRLIGEAFVDKPVSPWMVISPEAKINPCYBNACLPMFGKHKVLTQD 198

QY	120	LYSMKDDLWISIOTNGTKLATQMRKTSNLRIGFAFDKPKVPSYMISSPERALENPYD	179	A; Accession: S12534
Db	144	LSYSMKDDLKIQTGTLGTSLSERMRRLTSNLRIGFAFDKPKVPSYMISSPERALENPYD	203	A; Molecule type: mRNA
QY	180	MKTCLPMGGYKHVLTIDCQVTRNEEKKQSYRSRNRAPEGFDAMQATDEKIGWR	239	A; Residues: 1-644, 'L', 646-799 <RAM>
Db	204	FNTECMPTFGYKHVLTIDCQVTRNEEKKQSYRSRNRAPEGFDAMQATDEKIGWR	263	A; Cross-references: 1:644-1:646, 1:647-799
QY	240	NDASHLIVTIDAKTHIALDGRLAGIVQPDGQCHVGSDNHSASTMDYPSIGLMEKL	299	A; Map Position: 17q11-17qter
Db	264	NESSLLVLTIDAKTHIALDGRLAGIVQPDGQCHVGSDNHSASTMDYPSIGLMEKL	323	A; Superfamily: integrin beta chain; laminin-type EGF-like homology
QY	300	SOKNINLFAVTENVNLQYQNSYSELIGTIVGVLMSDSSNVQLIVDAYGKRSKVELEV	359	A; Keywords: cell adhesion; cytoskeleton; duplication; extracellular matrix; glycoprotein
Db	324	SDKNINLFAVTENVNLQYQNSYSELIGTIVGVLMSDSSNVQLIVDAYGKRSKVELEV	383	F:1-24/Domain: signal sequence #status predicted <SIG>
QY	360	RDLEFEELSFISFNATCLNNNEVPIGKSCMGKIGDVTSSIEAKVGRQPEKKSFTKPV	419	F:25-79/Product: integrin beta-5 chain #status predicted <EXT>
Db	384	RDLPEELSLFSASCONDELTPGLKSTGKIGDVTVERSEAKVRECPVSRQKTFKPV	443	F:720-794/Domain: laminin-type EGF-like homology <LEG>
QY	420	GFKSILIVOTFDCDCAQAOAQPNSHNCNGTFCGGVRCRGPMGLGQSCSEEDYR	479	F:463-513/Domain: transmembrane #status predicted <TM>
Db	444	II	503	F:743-799/Domain: intracellular #status predicted <INT>
QY	480	PSQDECSRREGQPVCSORGECLGQCVCHSSDFGKIGKCYCCEDDFSCVRKGEMSGH	539	F:347,460,477,505,552,586,654,705/Binding site: carbohydrate (Asn) #status
Db	504	PSQODRCSSKEGABVCSRRGECVGQCVGRASDGLKGWKGKCYCDDFSCLRYKGEMSGH	563	predicted <EXT>
QY	540	GQCSGDCCLSDMTGTYCNCCTRTDTOMSNSLLCSRGKRCRGSCVCIQPSGYDICE	599	Q: 1 GPNICITTRGVSOCIAVSPNCAMSDEALPLSPR----RCOLKENILKDCAPESE 55
Db	564	GQCNQGDCCESTGQDCNCITRTDTOMSNSLLCSRGQCVGKFCETQPSGYGTC	623	Q: 24 GLNCTICSGSATSCCECLILHPRKCAWCKE--DFGSPRSITSRCDLRLANLVNGGGE-IE 80
QY	600	KCPKCPDAGTCFKCCKVECKKFDKLADENTCNRYCDEIEVSKELDGTQAVNCTYKN	659	Q: 56 FPVSEARVLERPLSKGKGSDSS-QYTVQSPORTALRRLRDDSNSKNF1QVRQVEDYDPDI 114
Db	624	KCPCPDPACTFKKCCVCKKFERQPFWEFDSKQICROTEESTVYELDNGKAVNCTYKD	683	Q: 81 SPASSEFVRLSPLSKGSIGSAGWDVQDQIPOETAVLNRLQGDKTFPOLOVQRDYDPDI 140
QY	660	EDDCWVRCYQYEDSSGKSLILVYVEEPECPKGPDIWVLSSMGAIIIGLALLIWKLI	719	Q: 115 VYLMDSYSKDDLWISIQLNGLTKLATQMRKTSNLRIGFAFDKPKVPSYMISSPERALE 174
Db	684	ENDCVTFRQYHEDASGKSVLYVINEACPHGPDIWVLMVMSMGAIIIGLALLIWKLI	743	Q: 141 YYLMDLISLMSKMDLNLNTRLSLGTQKLAEMRKLTSNRLGIGSFVUDPSY-TAPRQ 199
QY	720	TIDRKERFAKFEERARKWDTANNPPLYKATSTFTNTYRG	761	Q: 175 NPC-YTMKTCPLPMGKYKHVLTIDCQVTRNEEKKQSYRSRNRAPEGFDAMQATWYC 232
Db	744	TIDRREFAKFEERARKWDTAHNPLYKGAATSTFTNTYRG	785	Q: 200 NPCIGKLFNPCKVPSFGFRHLPLTDVDSFNEVRQVRQVSRNRAPEGFDAVLQAVC 259
RESULT 9				
A; Accession: A38308		C; Species: Homo sapiens (man)		Q: 233 DEKIGRNFDASHLIVTIDAKTHIALDGRLAGIVQPDGQCHVGSDNHYASSTMDYSL 292
A; Accession: A38308		C; Date: 14-Jun-1991 #sequence_revision 14-Jun-1991 #text_change 20-Aug-1999		Q: 260 KEKIGRNDAHLVETTDVPHALDGGIGLWOPHDGQCHMNEETASNMDYSL 319
A; Accession: A38308; MUID:91009141		C; Cross-references: GB:J05633; NID:9186504; PIDN:AA59183.1; PID:9186505		Q: 293 GLMTEKISQKNINLFAVTENVNLQYQNSYSELIGTIVGVLMSDSSNVQLIVDAYGKIR 352
A; Accession: A38308		C; Note: parts of this sequence, including the amino end of the mature protein, were confirmed by cross-references		Q: 320 ALIGEKAENNLNLIFAVTRKHYHMLYKNTFALIRGTTVEILDGSKNITIOLINAYNSIR 379
A; Residues: 1-799 <MCL>		C; Residues: 1-799 <MCL>		Q: 353 SKVLEVRDLEFEELSFISFNATCLNNNEVPIGKSCMGKIGDVTWSIYEKVRGCP-QRKE 411
A; Cross-references: GB:J05633; NID:9186504; PIDN:AA59183.1; PID:9186505		C; Cross-references: GB:J05633; NID:9186504; PIDN:AA59183.1; PID:9186505		Q: 380 SKVELSWDQDPEDLNUFIFTACQDVSYPQQRKCEGLGDTASFEVLEARCSRSR 439
A; Residues: 1-799 <MCL>		C; Residues: 1-799 <MCL>		Q: 412 KSFTIKEVGFDLSLIVTIDCDCACQADQAREPNSHRCNGTFCGGVRCRGPGWLGS 471
A; Note: parts of this sequence, including the amino end of the mature protein, were confirmed by cross-references		C; Note: parts of this sequence, including the amino end of the mature protein, were confirmed by cross-references		Q: 440 HVFALKEVGRFDSLEVGVTNTCTCGCSVGLPEPNASC-NGSGTYVCGGECSPGVLGR 498
A; Cross-references: GB:J05633; NID:9186504; PIDN:AA59183.1; PID:9186505		C; Cross-references: GB:J05633; NID:9186504; PIDN:AA59183.1; PID:9186505		Q: 472 ECSEEDYRPSQDCEPSREGQPVCSORGECLGQCVCHSSDFGKIGKCYCCEDDFSCRY 531
A; Residues: 1-799 <MCL>		C; Residues: 1-799 <MCL>		Q: 499 ECQDGENQSYVQNLCREAGKPLCSGRGDCSCNCQSCSEFEGKLYGPCECDNFSCARN 558
A; Title: Cloning of an integrin beta subunit exhibiting high homology with integrin beta		C; Title: Cloning of an integrin beta subunit exhibiting high homology with integrin beta		Q: 532 KGMCSGGHGGQCGCGDCLSDMTGTYCNCCTRTDTOMSNSLLCSGRGKCECGSCVCTQ 591
A; Reference number: A35775; MUID:90319111		C; Reference number: A35775; MUID:90319111		Q: 559 KGVLCGSGHGGCHCGDCKCHAGYIGDNCNGSTDISTCRGQDQOSERHCLCQCGCTEP 618
A; Accession: A35775		C; Accession: A35775		Q: 592 GSYGDPREKECTCPDACTKFCVECKKFDKLADENTCNRYCDEIEVSKELDGTQ 650
A; Molecule type: mRNA		A; Molecule type: mRNA		Q: 619 GAFGENECKCPTCPDACCSTDRDCVCPCLHHSKGK-PDNQTCHSICRDEVITWVDTQDQ 677
A; Residues: 1-192, 'A', 194-644, 'L', 646-789, 793-799 <SUZ>		A; Residues: 1-192, 'A', 194-644, 'L', 646-789, 793-799 <SUZ>		Q: 651 DAVNCYKNEDDCIVRQYEDSSGKSLILVYVWEPCKGKPDILWLSVGMGILLGIA 710
A; Cross-references: GB:M35011; NID:9184524; PIDN:AA52707.1; PID:9306894		R; Romaswamy, H.; Hemler, M.E.		Q: 678 EAVLCFVYKTAKDCVAMFTVLPSPGSKNLNLVRLREPECGNTPNAMTILLAVGSSILVGLA 737
A; Title: Cloning, primary structure and properties of a novel human integrin beta subunit		A; Title: Cloning, primary structure and properties of a novel human integrin beta subunit		Q: 711 ALIWKLLITHDRKEFKEERARKWDTANNPPLYKATSTFT-----NITVFGT 762

Db	738 LLNIKLLVTDHRRERAKFOSESRARYEMASNPYRKPISTHTVDFENFKNSYNGT	797	Db	432 HIIKPKVGLGDALELLVSPCNCDCQKEVVEVNSKCHHNGSFOCGYCACHPGHMGPRCE	491
RESULT	10		Qy	473 CSEEDYRPSQDECSPREQQPVESQRGECLCGOCVCHSSDFGKTTGKYCECUDFSCVRYK	532
A37057	Integrin beta-6 chain - human		Qy	533 GEMCGSGHGSQCGSCGCLCSDWGGYCYCNGTTRDTCMSSNGLUGSGRCKCCEGSGCVCTPG	592
C;Species: Homo sapiens (man)	C;Date: 15-Feb-1991 #sequence_revision 13-Sep-1991 #text_change 19-Jan-2001		Db	492 CGED--MSTDSCKEAPDHPSGSGRCGCGCICLHSPPGNYIGYCOCONFCVYRK	548
C;Accession: A37057; 169201	R;Sheppard, D.; Rozzo, C.; Starr, L.; Quaranta, V.; Erie, D.J.; Pytel, R.		Qy	549 GLLCCGGNGDODCGECVCRSWGWTGBCYCNCITSTDSCVSEBDGVLCSGRGDCVCVGCVCTNPG	608
J; Biol. Chem. 265: 11502-11507, 1990	A;Title: Complete amino acid sequence of a novel integrin beta subunit (beta6) identified		Qy	593 SYGTCCEKCPCPDACTKKECKKEDRGGALHDENTCNRCDTESVKELKDQKD-	651
A;Reference number: A37057; MUID:90307659	A;Status: preliminary		Db	609 ASGPTCERCPCTGDPNCNSKRSCTECHLISAGQGEE--CVDKKLAGATISSEEDFSKD	666
A;Accession: A37057	A;Molecule type: mRNA		Qy	652 AVNCYTKNEDCVCVRFQYEDSSGKSLIVVERECPKGPPDILVLLSVMMGAILLGLAA	711
A;Residues: 1-788	A;Title: The gene organization of the human beta 7 subunit, the common beta subunit of the		Db	667 SVSOSLQGENECLITFLITTDNGKTIIHSINSKDCPQPPNPMMICGVSLATLIGWL	726
A;Reference number: 154749; MUID:93002753	A;Accession: 169201		Qy	712 LLNIKLLVTDHRRERAKFOSESRARYEMASNPYRKPISTHTVDFENFKNSYNGT	760
A;Status: preliminary; translated from GB/EMBL/DDBJ	A;Molecule type: DNA		Db	727 LCIWKLVLNSFHDKRKEVAKFEAERSKAKWQTGTPLYRGSTSFKNVYK	775
A;Residues: 1-157, 'R', 159-197 <JTA>	A;Cross-references: GB:549380; NID:9257588; PIDN:AAB23690.1; PID:9257589		Qy	777	
C;Genetics: GDB:ITGB6	A;Cross-references: GDB:131392; OMIM:147558		Db	778	
A;Map position: 2pter>qter	C;Superfamily: integrin beta chain; laminin-type EGF-like homology		Qy	779	
C;Keywords: blocked amino end; cell adhesion; cytoskeleton; glycoprotein; lipoprotein; membrane	F7/108-730/Domain: transmembrane #status predicted <TR>		Db	780	
F7/2/Modified site: myristylated amino end (GLY) (in mature form) #status predicted	F7/7/Binding site: palmitate (C16) (covalent) #status predicted		Qy	781	
F7/16,48,97,260,387,396,463,471/Binding site: carbohydrate (Asn) (covalent) #status predicted	Q: Best Local Similarity 51.9%; Score 2156; DB 2; Length 788; Matches 389; Conservative 50.6%; Pred. No. 1. 9e-129; Mismatches 231; Indels 16; Gaps 8;	RESULT 11	Db	782	
Query Match 1 GPNICTR----GVSSCQCLAVASPNCACSDDEAL--PLG-SPRCDLKENLKKNCAPE 52	Query Match 1 GPNICTR----GVSSCQCLAVASPNCACSDDEAL--PLG-SPRCDLKENLKKNCAPE 52	JC2005	Qy	783	
Db 14 GRNDSRTRWLCLGGAETCDEGLLIGPQCAWCQENFTHPSPGVGERCDTPANLLAKGCQLN 73	Db 14 GRNDSRTRWLCLGGAETCDEGLLIGPQCAWCQENFTHPSPGVGERCDTPANLLAKGCQLN 73	R;Shoji, N.; Hayzer, D.J.; Kim, T.M.; Runge, M.S.; Hanson, S.R.	Db	784	
Qy 53 SIEPVSEARVIEDRPLDKSGDSQVTVSPSPQFALRURPDDSKNFQVQRQVEDPYV 112	Qy 53 SIEPVSEARVIEDRPLDKSGDSQVTVSPSPQFALRURPDDSKNFQVQRQVEDPYV 112	Gene 133: 307-308, 1993	Qy	785	
Db 74 F1ENPVSOVETNLKPLSPLVGKRNQKNSDIVOIAPOQSLKLRPGGAQTLQHYRQDTPYV 133	Db 74 F1ENPVSOVETNLKPLSPLVGKRNQKNSDIVOIAPOQSLKLRPGGAQTLQHYRQDTPYV 133	A;Title: Human and baboon integrin beta 5 subunit-encoding mRNAs have alternative poly	Db	786	
Qy 113 DTYLMDLSYKDDWISQNLGKTLATOMRKLTSNLRIGFADFVDRKPVSPYPMIYSPPEALENPC 172	Qy 113 DTYLMDLSYKDDWISQNLGKTLATOMRKLTSNLRIGFADFVDRKPVSPYPMIYSPPEALENPC 172	A;Reference number: JC2005; MUID:94040831	Qy	787	
Db 134 DTYLMDLSAENDDLNTIKEGSGLSKESKTSNFRFLGFGFVEKPVSPFVK-TPEE 192	Db 134 DTYLMDLSAENDDLNTIKEGSGLSKESKTSNFRFLGFGFVEKPVSPFVK-TPEE 192	A;Accession: JC2005	Db	788	
Qy 173 LENPCYOMKTKTCLPMGQKYHVLTLTPOVTRENEVKKYKOSVRNDRAPEGGFDAMQATWC	Qy 173 LENPCYOMKTKTCLPMGQKYHVLTLTPOVTRENEVKKYKOSVRNDRAPEGGFDAMQATWC	A;Molecule type: mRNA	Qy	789	
Db 193 IANPCSSIPYFCPLPFGKHLPLNDAERENETIVKNOKISANIDTPEGGFDATMQAACV 252	Db 193 IANPCSSIPYFCPLPFGKHLPLNDAERENETIVKNOKISANIDTPEGGFDATMQAACV 252	A;Residues: 1-656 <SHO>	Db	790	
Qy 233 DEKIGRNDAHLVLTDAKTHALDGRLAGIVOPNDGCHQVSYSDNHYASSTDMDYPL	Qy 233 DEKIGRNDAHLVLTDAKTHALDGRLAGIVOPNDGCHQVSYSDNHYASSTDMDYPL	C;Note: the authors translated the codon AGA for residue 454 as Lys. CAA for residue	Qy	791	
Db 253 KERIGWRNDLHLHLVFDSDHFGMSKLAGIVIPDGLCHDSKNEYSMSTVLEYPT	Db 253 KERIGWRNDLHLHLVFDSDHFGMSKLAGIVIPDGLCHDSKNEYSMSTVLEYPT	C;Comment: This protein is a predominant subunit for the vitronectin receptor in babo	Db	792	
Qy 293 GLMTEKLSQKINLIFAVTENVNLQYNSLIPGTVGVLMSDSSNVQLOLIVDAYKIR 352	Qy 293 GLMTEKLSQKINLIFAVTENVNLQYNSLIPGTVGVLMSDSSNVQLOLIVDAYKIR 352	C;superfamily: integrin beta chain; laminin-type EGF-like homology	Qy	793	
Db 313 GQDILKLVVNNLLIAVTOPOQHENVYAKLICATVGLOKQSGNLIQLLSAYEER 372	Db 313 GQDILKLVVNNLLIAVTOPOQHENVYAKLICATVGLOKQSGNLIQLLSAYEER 372	C;Keywords: cell adhesion	Db	794	
Qy 353 SKVLEVLRDPEELSLFSNATCNNEVPLGKSCMGLKICDVTWSIEAKVRGCP-OEKSF	Qy 353 SKVLEVLRDPEELSLFSNATCNNEVPLGKSCMGLKICDVTWSIEAKVRGCP-OEKSF	F352	Qy	795	
Db 373 SEVLEVLGQDPEGLNSFTAICNGTTLFQHOKKCSHMKVDTASFSVTWIPHC-ERRS	Db 373 SEVLEVLGQDPEGLNSFTAICNGTTLFQHOKKCSHMKVDTASFSVTWIPHC-ERRS	431	Db	796	
Qy 413 SFTIKPGFKSLLIVOTFDCDCACQQAQABENSHCNCNGTFCGGVCRGPGWLGSOCE	Qy 413 SFTIKPGFKSLLIVOTFDCDCACQQAQABENSHCNCNGTFCGGVCRGPGWLGSOCE	472	Qy	797	

Db 359 DGEHHSVYQNLCRTEGKLCGSGDCSCNQCSCEFSFGKLYIGPFCFCDCDNASCARNKGV 418
 Qy 535 MCGHIGOCSCGDCCLCSDMVGYCNCTTRDTGCMSSNGLLCSRGKCBGSCVCTGGSY 594
 Db 419 LCGSGHGECHGECKHAGYIGDNCNCSTIDTICRGDRGQICSERGHCILCGOCCTEGAF 478
 Qy 595 GDTCEKCPICPDPACTEKKCCKEKKFEDGALHQDENTCNYCROIES-VKELKDTGDAV 653
 Db 479 GEMCBEKCPICPDCATSKRDCVECLPLHSGK-PDNOTCSLCRDEVITWDITVKDDOEV 537
 Qy 654 NCTVKNEDCQVRRQYEDSGKSILYVVEPECPKGDILVLLSIVMAGILLIGLA 713
 Db 538 LCFVTKTAKOCVMMETYVELPSGSNLTULREPECGNPNTNAMTILLAVGSLVLLA 597
 Qy 714 IWKLLITIDRKEFKEFEEFEEARAKWDTANNPYKEATSTFT-----NITYRGT 762
 Db 598 IWKLLITIDRKEFKEFEEFEEARAKWDTANNPYKEATSTFT-----NITYRGT 654

RESULT 12

JGCH3
 Integrin, band 3 precursor - chicken
 N;Alternate names: CSAT antigen; JG22 antigen; RGD-receptor
 C;Species: Gallus gallus (chicken)
 C;Date: 31-Mar-1988 #sequence_revision 31-Mar-1988 #text_change 23-Jul-1999
 C;Accession: A23947
 R;Tankun, J.W.; Desimone, D.W.; Fonda, D.; Patel, R.S.; Buck, C.; Horwitz, A.F.; Hynes, Cell, 46, 271-282, 1986
 A;Title: Structure of integrin, a glycoprotein involved in the transmembrane linkage between
 A;Reference number: A23947; MUID:86245073
 A;Accession: A23947
 A;Molecule type: mRNA
 A;Residues: 1-803 <PAM>
 A;Cross-references: GB: M14049; NID:9212213; PIDN:AA48926.1; PID:9212214
 A;Experimental source: embryonic fibroblasts
 A;Note: the amino end of the mature protein is blocked
 C;Comment: This transmembrane complex may be the target of oncogenic transformation that
 C;Comment: The extracellular domain of this subunit contains a potential tyrosine-kinase phosphorylation site.
 C;Keywords: cell adhesion; cytoskeleton; duplication; extracellular matrix; glycoprotein
 F;1-24/Domain: signal sequence #status predicted <SIG>
 F;25-803/Product: integrin, band 3 #status predicted <MAT>
 F;25-73/Domain: extracellular #status predicted <EXI>
 F;467-654/Region: cysteine-rich
 F;467-654/Region: duplication
 F;734-755/Domain: transmembrane #status predicted <MEM>
 F;757-803/Domain: intracellular #status predicted <INP>
 F;216-273, 367-410, 421-433, 445, 486-525, 589-624, 674/Binding site: carbohydrate (Asn) (cova)
 F;788/Binding site: phosphate (Tyr) (covalent) #status predicted

Query Match
 Best Local Similarity 44.5%; Score 1846.5; DB 1; Length 803;
 Matches 353; Conservative 136; Mismatches 226; Indels 27; Gaps 15;

Db 526 FSCVRYKGMGSGHGCSCGCGCLCSDWTGKYCNCNTTDTGCMSSNGLLCSRGKCGCGS 585
 Db 563 FNCDSRNGLIGGNGKCRVCECIPNFTGSAACDSDTTPCMAGNGQTCNGRCBEGT 622
 Qy 586 CYC1QPSYGDPCPDPACTEKKCCKEKKFEDGALHQDENTCNYCROIES-VKELKDTGDAV 641
 Db 623 CNCTDPKFQFGPCMCQTCGLGCAEHKDCVQCRAFEGK-EKKEICSCBHMNTRV 680
 Qy 642 VKEI-KTGKAV-NOTYKNEEDCQVRRQYEDSGKSILYVVEPECPKGDILVLLS 699
 Db 681 RGLKQPOVHPDPLSHCREKDVGDCWPFYFTSVNSNGEASVHVPETCPKGDPDITPIVAG 740
 Qy 700 VMGAILLIGLALLWMLLITHDRKEFKEFEEFEEARAKWDTANNPYKEATSTFTNITY 759
 Db 741 VVAGIVVILGALLWMLLITHDRKEFKEFEEFEEARAKWDTANNPYKEATSTFTNITY 800
 Db 801 EG 802

RESULT 13

B27079
 fibronectin receptor beta chain precursor - human
 N;Alternate names: CD29 antigen; integrin beta-1 chain
 C;Species: Homo sapiens (man)
 C;Date: 30-Jun-1988 #sequence_revision 30-Jun-1988 #text_change 20-Aug-1999
 C;Accession: B27079
 R;Argravies, W.S.; Suzuki, S.; Arai, H.; Thompson, K.; Pierschbacher, M.D.; Ruoslahti, J. Cell Biol, 105, 1183-1190, 1987
 A;Title: Amino acid sequence of the human fibronectin receptor.
 A;Reference number: A27079; MUID:88007843
 A;Accession: B27079
 A;Molecule type: mRNA
 A;Residues: 1-198 <ARG>
 A;Cross-references: EMBL:X07979; NID:931441; PIDN:CAA30790.1; PID:931442
 C;Genetics:
 A;Gene: GDB:1TGB1; FNFB
 A;Cross-references: GDB:118732; OMIM:135630
 A;Map Position: 10p11.2-10p11.2
 C;Superfamily: integrin beta chain; laminin-type EGF-like homology
 C;Keywords: alternative splicing; duplication; heterodimer; transmembrane protein
 F;1-20/Domain: signal sequence #status predicted <SIG>
 F;21-798/Product: fibronectin receptor beta chain #status predicted <MAT>

Query Match
 Best Local Similarity 44.1%; Score 1831; DB 2; Length 798;
 Matches 350; Conservative 134; Mismatches 268; Indels 28; Gaps 14;

Db 147 IDLVYMLDLSYSSMDKDLRKNLKSLSGKMEKTSFRRIGGSFVKTWYISTTPAK 206
 Qy 172 ALENFCYDMKTTCLPMFGYKHLVLTQDVTRENEEVKQSVRNDRPEEGFADIMAV 231
 Db 87 PRGSKRVLDEREVTRNKGAAEKLPEAITQIOPQKVLQLRVRGEPOTPSLKFRAEDY 146
 Qy 112 VDIYLMDSYSSMDKDLIISIQLNGLTKATQMRKLTNSNRIGGAFVDKPVPSVIMYISPE 171
 Db 147 IDLVYMLDLSYSSMDKDLRKNLKSLSGKMEKTSFRRIGGSFVKTWYISTTPAK 206
 Qy 172 ALENFCYDMKTTCLPMFGYKHLVLTQDVTRENEEVKQSVRNDRPEEGFADIMAV 231
 Db 207 -LRNPCTG-DQNTSPFSKVNVLSTISRGKNEFELVGQHISGNLSDPSEGFDAMOVAV 264
 Qy 3 NICTRGVSQCLAVSPMCACSD----EALPLGSPRCDEKLENLKDNCAPESTIFP 57
 Db 25 NRCLKANAKSCGECIOAGAPNCGWCTNSTLQEGMPT-SARCDLLEALKKKGCPDDINP 83

F1-20/Domain: signal sequence #status predicted <SIG>
F21-798/Product: integrin beta-1 chain #status predicted <MMT>

Query	Match	43.8%	Score	1821	DB	28	length	798
Best	Local	Similarity	44.4%	Pred.	No.	3	9e-18	
Matches	346	Conservative	137	Mismatches	269	Indels	28	Gaps
QY	3	NICUTRGYSSCOOLAYSPMCACSD	---EALPLGSPPRDLKENLKDNCAPSEIEFP	57				
Db	25	NRCIKNANAKSCGCICQAGPNCCWCTNTFLQGDMPT-SARDDELLAKKGCQSDIEPN	83					
QY	58	VSEARVLEDPLSDDKGSSDQ	---VTOVSPIRALRURPDDSKNFSTQVROVEDPY	112				
Db	84	RGSOTIKKNNVNTNSKGMAEKURPEDITOQQPQOLLKURSGEFOKFTLKKRKAEDPY	143					
QY	113	DIYLMQDLSYSHKDDLMWSIQNIGTKLATQMRKTSMLRTGFAFYDKPPSPYMTSPPA	172					
Db	144	DUJYLMQDLSYSHKDDLMENVKSLGTDLMNEMRRTSDPRIGFSFVEKTYMPYISTPAK-	202					
QY	173	LENPCYDMDTCTLPMLPGYKHLTLTDQVTRNEEYKQSVSRRNDAPEGFDAMQATC	232					
Db	203	LRNPPC-TSEQNQTSPPSYKVNLSLTDGRERENLWQORIGNSLUDSPEGGDAMQAVC	261					
QY	233	DEKIGARNDASHLVLVFTDAKTHIALDGRLAGIVQPNPDGQCHVGSDNHYSASTMDSL	292					
Db	262	GSLIGWNRN-VTRLYFESTDAGFHAGDGKGGIVLPNDQOCHL-ENVYVTHMHDYFSI	319					
QY	293	GLMTEKLSQKINLIFAVTENVNLYQNTYSELINGPTVGSLSMOSSTNVIQLTYDAYKIR	352					
Db	320	AHLVQKUSSENNIQTFAVTREHFQPYKELKLNLIPKSAVGLTSQGNSSNVNQQLIDAYNSL	379					
QY	353	SKEVLEVRDPEELLSFENATCLN-NEVPIGLSCMGKJGDSVSFEAKYRGCOEK	410					
Db	380	SEVILENSKLDLGTVNLYKSYCKNGVNGTGEENGKCRCSNISIGERVQEFETISANTCENKE	439					
QY	411	EKFSETIKPVGFKDSLIVQVTFDCPACQQAEPNSHRNCNNGTFECGYCROGPWGWSQ	470					
Db	440	SETIKIKPLGFTEEVWVLOFICKNCNSHGTASPCKHEGNGTFECGACRNEGRYGRH	499					
QY	471	CECSEEDYRSSQD-CSRREGQPVCSORGECILCGQCVCHSSDFRK-ITGKCECIDS	527					
Db	500	CECSATDEVNEDMDAAYCRKRNNSSEFICSNNGECVCGQCVRKRONTNEYTSGKFCCECDNF	559					
QY	528	CVRYKGEMCGSGHGOCSCGDCPCLCSDWDIGYCNTRTDTMMSNLICSGRKCECGCSV	587					
Db	560	CDRNSGLICGGNGVCRCRVCECIPNYSACDSDLNGPCLASNGQCGRCNGICEGACK	619					
QY	588	CIPCGSYGSGTCKECPCTPDACTFKCCEVCKEDFGALDENTCNRYCR---DETESVK	643					
Db	620	CTDPKFQGQPCETCQTCGLGVCAEHEKCVOCRAFNKGKD-TCAQZCSHNLNTKVES	677					
QY	644	ELKPTGK - DAWVCTYKNEDDCVWRQYIEDSGKSILYVVEPECEPKGPDILVILSV	701					
Db	678	KLPOVQPVWPTVCKEKKDIDWYFTSYVNGNEAVHVVETPDCPTGPDIPIVAGV	737					
QY	702	GAIILIGIALLWKLWKLITIHRKEFAKEFEEERARAKWDTANPLKAEATSTFTNTYRG	761					
Db	738	AGIVLIGIALLWKLWKLMIHORREFAKFEKKNAKWDTCENPISAVTWWPKY	797					

Search completed: May 19, 2002, 12:24:02
Job time: 9709 sec

Gencore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

Om protein - protein search, using sw model

Run on: May 19, 2002, 10:39:46 ; Search time 40.44 Seconds

(without alignments)
729.582 Million cell updates/secTitle: US-09-673-302a-1
Perfect score: 4154

Sequence: 1 GPNICTTREVSSCQCLAVS.....NNPPLYKEATSTFFNITYRGT 762

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Swissprot_40:*

Pred. No. 18 is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	411.9	99.2	788	1 ITB3_HUMAN
2	303.9	92.4	787	1 ITB3_MOUSE
3	234.3	798	1 ITB5_MOUSE	
4	233.9	56.3	799	1 ITB5_HUMAN
5	217.7	52.4	787	1 ITB6_MOUSE
6	210.5	52.3	788	1 ITB6_HUMAN
7	198.9	47.9	655	1 ITB5_PAPC
8	184.5	44.5	803	1 ITB1_CHICK
9	183.3	44.1	798	1 ITB1_FELCA
10	183.1	44.1	798	1 ITB1_HUMAN
11	182.5	44.0	798	1 ITB0_XENLA
12	182.1	43.8	798	1 ITB1_MOUSE
13	181.8	43.8	798	1 ITB1_XENLA
14	181.7	43.7	773	1 ITB1_BOVIN
15	180.1	43.4	799	1 ITB1_RAT
16	167.6	40.3	577	1 ITB6_CAVPO
17	157.9	38.0	809	1 ITB3_CAEEL
18	153.7	37.0	845	1 ITB8_DRONE
19	152.0	36.6	769	1 ITB2_BOVIN
20	151.5	36.5	771	1 ITB2_MOUSE
21	151.3	36.4	769	1 ITB2_HUMAN
22	150.9	36.3	769	1 ITB2_PIG
23	150.6	36.3	806	1 ITB7_HUMAN
24	150.2	36.2	798	1 ITB8_RABBIT
25	133.1	32.0	768	1 ITB8_HUMAN
26	132.6	31.9	1 ITB4_HUMAN	
27	118.4	28.5	1822	1 ITB4_HUMAN
28	115.4	27.8	1807	1 ITB4_RAT
29	26.4	7.9	1 ITB8_DRONE	
30	260.5	6.3	4289	1 TENK_HUMAN
31	25.9	6.2	2201	1 TENA_HUMAN
32	25.5	6.1	1746	1 TENA_PIG
33	252.5	6.1	1808	1 TENA_CHICK

RESULT	1	ALIGNMENTS
ID	ITB3_HUMAN	
ITB3_HUMAN STANDARD; PRT; 788 AA.		
AC	P05106; Q13413; Q16999; Q14648; Q15495;	
DT	13-AUG-1987 (Rel. 05, Created)	
DT	13-AUG-1987 (Rel. 05, Last sequence update)	
DT	01-MAR-2002 (Rel. 41, Last annotation update)	
DE	Integrin beta-3 precursor (Platelet membrane glycoprotein IIIa)	
DE	(GP1IIa) (CD61).	
GN	ITGB3 OR GP3A.	
OS	Homo sapiens (Human).	
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
OX	NCBI_Taxid=9606;	
RN	[1]	
RP	SEQUENCE FROM N.A. (ISOFORM BETA-3A).	
RX	MEDLINE=87165931; PubMed=3494014;	
RA	Frizzell L.A., Steinberg B., Rall S.C. Jr., Lo S., Phillips D.R., Prandini M.H.,	
RT	"Protein sequence of endothelial glycoprotein IIIa derived from a cDNA clone. Identity with platelet glycoprotein IIIa and similarity to 'integrin'."	
RT	J. Biol. Chem. 262:3936-3939 (1987).	
RL		
RN	[2]	
RP	SEQUENCE FROM N.A. (ISOFORM BETA-3A).	
RX	MEDLINE=90265363; PubMed=2315548;	
RA	Frachet P., Usan G., Thevenon D., Denarier E., Prandini M.H., Marquerie G.;	
RT	"GP1IIb and GP1IIa amino acid sequences deduced from human megakaryocyte cDNAs."	
RL	Mol. Biol. Rep. 14:27-33 (1990).	
RT	[3]	
RP	SEQUENCE FROM N.A. (ISOFORM BETA-3A).	
RX	MEDLINE=88213696; PubMed=2452834;	
RA	Zimrin A.B., Eisman R., Vilair G., Schwartz E., Bennett J.S., Ponce M.;	
RT	"Structure of platelet glycoprotein IIIa. A common subunit for two different membrane receptors."	
RT	J. Clin. Invest. 81:1470-1475 (1988).	
RL	[4]	
RP	SEQUENCE FROM N.A. (ISOFORM BETA-3A).	
RX	MEDLINE=9025678; PubMed=2341395;	
RA	Zimrin A.B., Gladwitz S., Lord S., Schwartz E., Bennett J.S., White G.C., II, Ponce M.;	
RT	"The genomic organization of platelet glycoprotein IIIa."	
RT	J. Biol. Chem. 265:8590-8595 (1990).	
RL	[5]	
RP	SEQUENCE FROM N.A. (ISOFORM BETA-3C).	
RC	SEQUENCE FROM N.A. (ISOFORM BETA-3C).	
RT	"Cloning and characterization of a novel integrin beta3 subunit."	
RT	J. Biol. Chem. 272:16390-16397 (1997).	
RL	[6]	
RP	SEQUENCE OF 1-26 FROM N.A.	

RC TISSUE-Blood;
 RX MEDLINE=94129007; PubMed=8298129;
 RA Villa-Garcia M., Li L., Rieley G., Bray P.F.;
 RT "Isolation and characterization of a TATA-less promoter for the human
 RT beta 3 integrin gene.";
 RL Blood 83:668-676(1994).
 RN [7] SEQUENCE OF 27-788 FROM N.A. (ISOFORM BETA-3A).
 RX MEDLINE=91009291; PubMed=2142280;
 RA Lanza F., Kieffer N., Phillips D.R., Fitzgerald L.A.;
 RT "Characterization of the human platelet glycoprotein IIIa gene.
 RT Comparison with the fibronectin receptor beta-subunit gene.";
 RL J. Biol. Chem. 265:18098-18103(1990).
 RN [8] SEQUENCE OF 122-204 FROM N.A.
 RX MEDLINE=9300275; PubMed=1382574;
 RA Jiang W.M., Jenkins D., Yuan Q., Leung E., Choo K.H., Watson J.D.,
 RA Krissansen G.W.;
 RT "The gene organization of the human beta 7 subunit, the common beta
 subunit of the leukocyte integrins HML-1 and LPAM-1.";
 RL Int. Immunol. 4:1031-1040(1992).
 RN [9] PARTIAL SEQUENCE FROM N.A. (ISOFORM BETA-3B).
 RX MEDLINE=89315807; PubMed=2787511;
 RA Van Kuppevel T.H.M.S.M., Languino L.R., Gailit J.O., Suzuki S.,
 RA Ruoslahti E.;
 RT "An alternative cytoplasmic domain of the integrin beta 3 subunit.";
 RL PROC. Natl. Acad. Sci. U.S.A. 86:5415-5418(1989).
 RN [10] SEQUENCE OF 218-234 AND 439-443.
 RX MEDLINE=8710150; PubMed=3801670;
 RA Hiraiwa A., Matukage A., Shiko H., Takahashi T., Naito K., Yamada K.;
 RT "Purification and partial amino acid sequence of human platelet
 membrane glycoproteins IIb and IIIa.";
 RL Blood 69:560-564(1987).
 RN [11] PARTIAL SEQUENCE AND DISULFIDE BONDS.
 RX MEDLINE=91158732; PubMed=2001252;
 RA Calvete J.J., Henschel A., Gonzalez-Rodriguez J.;
 RT "Assignment of disulfide bonds in human platelet GPIIIa. A
 RT disulphide pattern for the beta-subunits of the integrin family.";
 RL Biochem. J. 274:63-71(1991).
 RN [12] PHOSPHORYLATION AT TYR-773 AND TYR-785 (ISOFORM BETA-3A).
 RX MEDLINE=96210016; PubMed=6631894;
 RA Law D.A., Nannizzi-Alaimo L., Phillips D.R.;
 RT "Outside-in integrin signal transduction. Alpha IIb beta 3 (GP IIb
 IIIa) tyrosine phosphorylation induced by platelet aggregation.";
 RL J. Biol. Chem. 271:10811-10815(1996).
 RN [13] X-RAY CRYSTALLOGRAPHY (3.1 ANGSTROMS) OF 27-78.
 RX MEDLINE=21482770; PubMed=1546339;
 RA Xiong J.-P., Stetler T., Dieffenbach B., Zhang R., Dunker R., Scott D.L.,
 RA Joachimiak A., Goodman S.L., Arnout M.A.;
 RT "Crystal structure of the extracellular segment of integrin alpha
 RT Vbeta3";
 RL Science 294:339-345(2001).
 RN [14] VARIANT HPA-1 (PLA1).
 RX MEDLINE=89214713; PubMed=2565345;
 RA Newman P.J., Derbes R.S., Aster R.H.;
 RT "The human platelet alloantigens, PLA1 and PLA2, are associated with
 RT a leucine33/proline33 amino acid polymorphism in membrane
 RT glycoprotein IIIa, and are distinguishable by DNA typing.";
 RL J. Clin. Invest. 83:1778-1781(1989).
 RN [15] VARIANT HPA-4 (PEN).
 RX MEDLINE=93035444; PubMed=1430225;
 RA Wang R., Furukata K., McFarland J.G., Friedman K., Aster R.H.,
 RA Newman P.J.;
 RT "An amino acid polymorphism within the RGD binding domain of platelet
 RT membrane glycoprotein IIIa is responsible for the formation of the

RT Peña/Pemb alloantigen system.";
 RL J. Clin. Invest. 90:2038-2043(1992).
 RN [16] VARIANT MO-1.
 RX MEDLINE=93112977; PubMed=8093349;
 RA Kuijpers R.W.A.M., Simsek S., Faber N.M., Goldschmeding R.,
 RA van Wermekken R.K.V., von Dem Bonne A.E.G.K.;
 RT "Single point mutation in human glycoprotein IIIa is associated with
 RT a new platelet-specific alloantigen (Mo) involved in neonatal
 RT alloimmune thrombocytopenia.";
 RL Blood 81:70-76(1993).
 RN [17] VARIANT CA/TU.
 RX MEDLINE=94063373; PubMed=7694683;
 RA Wang R., McFarland J.G., Kekomaki R., Newman P.J.;
 RT "Amino acid 489 is encoded by a mutational 'hot spot' on the beta 3
 RT integrin chain: the C-terminal human platelet alloantigen system.";
 RL Blood 82:3380-3391(1993).
 RN [18] VARIANT SR(A).
 RX MEDLINE=9417229; PubMed=8132570;
 RA Santoso S., Kalb R., Kroll H., Walka M., Kieffel V.,
 RA Mueller-Eckhardt C., Newman P.J.;
 RT "A point mutation leads to an unpaired cysteine residue and a
 RT molecular weight polymorphism of a functional platelet beta 3 integrin
 RT subunit. The Sra alloantigen system of GPIIIa.";
 RL J. Biol. Chem. 269:8439-8444(1994).
 RN [19] VARIANT GTPA TYR-145.
 RX MEDLINE=90364410; PubMed=2392682;
 RA Loftus J.C., O'Toole T.E., Plow E.F., Glass A., Frelinger A.L. III,
 RA Ginsberg M.H.;
 RT "A beta 3 integrin mutation abolishes ligand binding and alters
 RT divalent cation-dependent conformation.";
 RL Science 249:915-918(1990).
 RN [20] VARIANT GLN-240.
 RX MEDLINE=92156115; PubMed=1371279;
 RA Bajt M.L., Ginsberg M.H., Frelinger A.L. III, Berndt M.C.,
 RA Loftus J.C.;
 RT "A spontaneous mutation of integrin alpha IIb beta 3 (platelet
 RT glycoprotein IIb-IIIa) helps define a ligand binding site.";
 RL J. Biol. Chem. 267:3780-3794(1992).
 RN [21] VARIANT TRP-240.
 RX MEDLINE=92291320; PubMed=1602006;
 RA Lanza F., Stetler A., Fournier D., Morales M., Andre G., Nurden A.T.,
 RA Cazenave J.-P.;
 RT "A new variant of Glanzmann's thrombasthenia (Strasbourg T).
 RT Platelets with functionally defective glycoprotein IIb-IIIa complexes
 RT and a glycoprotein IIIa 214Arg->214Trp mutation.";
 RL J. Clin. Invest. 89:1995-2004(1992).
 RN [22] VARIANT STRASBOURG-1.
 RX MEDLINE=9306201; PubMed=1438206;
 RA Chen Y.-P., Djaffar T., Pidard D., Steiner B., Cieutat A.-M.,
 RA Caen J.P., Rosa J.-P.;
 RT "Ser-752->Pro mutation in the cytoplasmic domain of integrin beta 3
 RT subunit and defective activation of platelet integrin alpha IIb beta
 RT 3 (glycoprotein IIb-IIIa) in a variant of Glanzmann thrombasthenia.";
 RL Proc. Natl. Acad. Sci. U.S.A. 89:10165-10173(1992).
 RN [23] REVIEW ON GTPA VARIANTS.
 RX MEDLINE=95184171; PubMed=7878622;
 RA Bray P.F.;
 RT "Inherited diseases of platelet glycoproteins: considerations for
 RT rapid molecular characterization.";
 RL Thromb. Haemost. 72:492-502(1994).
 RN [24] VARIANT GTPA TRP-143.
 RX MEDLINE=98025992; PubMed=9376589;
 RA Basani R.P., Brown D.L., Vilaine G., Bennett J.S., Poncz M.;
 RT "A Leu17->Trp mutation within the RGD-peptide cross-linking region

Query Match	Score 4119;	DB 1;	Length 788;
Best Local Similarity	99.2%	DB 1;	Length 788;
Matches	757;	Conservative	0;
Matches	757;	Mismatches	5;
Indels	0;	Gaps	0;
STRAIN=C3H/HEN;		RC	RA
McHugh K. P., Teitelbaum S. L., Kitazawa S., Ross F. P.;		RA	RL
Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.		RL	CC
-1- FUNCTION: INTEGRIN ALPHA-V/BETA-3 IS A RECEPTOR FOR CYTOSACTIN, FIBRONECTIN, LAMININ, MATRIX METALLOPROTEINASE 2, OSTEOPONTIN, PROTHROMBIN, THROMBOSPONDIN, VITROECTIN AND VON WILLEBRAND FACTOR. INTEGRIN ALPHA-TIB/BETA-3 IS A RECEPTOR FOR FIBRONECTIN, FIBRINOGEN, PLASMINOGEN, PROTHROMBIN, THROMBOSPONDIN AND VITROECTIN. INTEGRINS ALPHA-TIB/BETA-3 AND ALPHA-V/BETA-3 RECOGNIZE THE SEQUENCE R-G-D IN A WIDE ARRAY OF LIGANDS. INTEGRIN ALPHA-TIB/BETA-3 RECOGNIZES THE SEQUENCE H-H-L-G-A-K-O-A-G-D-V IN FIBRINOGEN GAMMA CHAIN. FOLLOWING ACTIVATION INTEGRIN ALPHA-TIB/BETA-3 BRINGS ABOUT PLATELET/PLATELET INTERACTION THROUGH BINDING OF SOLUBLE FIBRINOGEN. THIS STEP LEADS TO RAPID PLATELET AGGREGATION WHICH PHYSICALLY PLUGS RUPURED ENDOTHELIAL SURFACE.	CC	CC	
1 GRNICTATRGVSSCQCLAVSPVCAWCSDDEALPLGSPRDLKENLKDNCAPSSIEFPVSE 60	1	CC	CC
27 GPNCIICATKGVSSCQCLAVSPVCAWCSDDEALPLGSPRDLKENLKDNCAPSSIEFPVSE 86	27	CC	CC
61 ARVLEDPLSDKGSDQSVQTVQSPQRALRPPDSSKNSVQRQVEDYPDIYMLD 120	61	CC	CC
87 ARVLEDPLSDKGSDQSVQTVQSPQRALRPPDSSKNSVQRQVEDYPDIYMLD 146	87	CC	CC
121 SSMKDIWISIQLNGTKLATORMLKTSMLRIGAFVQDVKPVSMYISPEALENPYCDM 180	121	CC	CC
207 KTCIPLMPGKVKVLTQDVTQFNEEVKKQKOSRSRNRAPEGGDAIMQATVDEKIGRN 266	207	CC	CC
241 DASHLYVFTDKTHIALDGRAGIVOPNDQCHVGSDNHYSATTMOPSYLGMTEKLS 300	241	CC	CC
267 DASHLYVFTDKTHIALDGRAGIVOPNDQCHVGSDNHYSATTMOPSYLGMTEKLS 326	267	CC	CC
301 QRMNLLFAVTENVNLLQNYSLIPGTVGISMDSNNVQLVDAIGRINKSVELEVR 360	301	CC	CC
327 QRMNLLFAVTENVNLLQNYSLIPGTVGISMDSNNVQLVDAIGRINKSVELEVR 386	327	CC	CC
361 DPEEELSLSENATCLNNHEVTPILKSCMGLKTQDVTSSIEAVKVRGCPQEKESTIKPGV 420	361	CC	CC
387 DLPEEELSLSFNATCLNNHEVTPILKSCMGLKTQDVTSSIEAVKVRGCPQEKESTIKPGV 446	387	CC	CC
421 FRLSLITQVTFCCDCAQQAEPNSHRCNGNTFECGVCRGPGWLGSCQCESEEDYRP 480	421	CC	CC
507 SQDECSPREGPVCSCORGECGCCGQCYCHSSPFGKTYGKCYCSCDDESCVRYGEMCGHG 540	507	CC	CC
447 FRLSLIVQVTFCCDCAQQAEPNSHRCNGNTFECGVCRGPGWLGSCQCESEYRP 506	447	CC	CC
541 QSCGDCGCDSDWYTGYCNCNTTRDGMSSGMSNLCSORGKCKCGSCVCIQPSYGDCEK 600	541	CC	CC
567 QSCGDCGCDSDWYTGYCNCNTTRDGMSSGMSNLCSORGKCKCGSCVCIQPSYGDCEK 626	567	CC	CC
601 CPTCPDACTFKECVECKFDGALHDENTCNYCROBIESKELKOTGKDAVNTKNE 660	601	CC	CC
627 CPTCPDACTFKECVECKFDPEPYMTENTCNYCROBIESKELKOTGKDAVNTKNE 686	627	CC	CC
661 DCVCRVQYEESSGKSLYVVEPECPKGPIVLUVLSVKMATTLLGLAALIWKLLT 720	661	CC	CC
687 DCVCRVQYEESSGKSLYVVEPECPKGPIVLUVLSVKMATTLLGLAALIWKLLT 746	687	CC	CC
721 IHRKEEAKFEERAKWDTANNPLYKEATSTFTNTYRG 762	721	CC	CC
747 IHRKEEAKFEERAKWDTANNPLYKEATSTFTNTYRG 788	747	CC	CC
RESULT 2			
BTB3_MOUSE	STANDARD;	PRM;	787 AA.
BTB3_MOUSE			
054890;			
15-DEC-1998 (Rel. 37, Created)			
15-DEC-1998 (Rel. 37, Last sequence update)			
01-MAR-2002 (Rel. 41, Last annotation update)			
Integrin beta-3 Precursor (platelet membrane glycoprotein IIIa) (GP11a) (CD61)			
ITGB3.			
Mus musculus (Mouse).			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
NCBI_TaxID=10090;			
11			

FT DISUFLID 482 520 BY SIMILARITY.
 FT DISUFLID 487 496 BY SIMILARITY.
 FT DISUFLID 498 511 BY SIMILARITY.
 FT DISUFLID 526 531 BY SIMILARITY.
 FT DISUFLID 528 561 BY SIMILARITY.
 FT DISUFLID 533 546 BY SIMILARITY.
 FT DISUFLID 548 553 BY SIMILARITY.
 FT DISUFLID 567 572 BY SIMILARITY.
 FT DISUFLID 569 600 BY SIMILARITY.
 FT DISUFLID 574 583 BY SIMILARITY.
 FT DISUFLID 585 592 BY SIMILARITY.
 FT DISUFLID 606 611 BY SIMILARITY.
 FT DISUFLID 608 656 BY SIMILARITY.
 FT DISUFLID 613 623 BY SIMILARITY.
 FT DISUFLID 626 629 BY SIMILARITY.
 FT DISUFLID 633 642 BY SIMILARITY.
 FT DISUFLID 639 712 BY SIMILARITY.
 FT DISUFLID 660 688 BY SIMILARITY.
 FT CARBOHYD 345 345 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 396 479 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 477 584 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 679 679 N-LINKED (GLCNAC. . .) (POTENTIAL).
 MOD. RES 772 772 PHOSPHORYLATION (BY SIMILARITY).
 FT MOD. RES 784 784 PHOSPHORYLATION (BY SIMILARITY).
 SQ SEQUENCE 86694 MW; B1570599ABC43BA3 CRC64;

Query Match 92.4%; Score 3839; DB 1; Length 787;
 Best Local Similarity 91.3%; Pred. No. 9.7e-253; Mismatches 27; Indels 0; Gaps 0;
 Matches 694; Conservative 39; Mismatches 27; Indels 0; Gaps 0;

QY 3 NICTRGRVSSCQOCLAVSPMCAMSCDEALPLGSPRCDIENILKDNCAPESTEPVSEAR 62
 Db 28 NICTRGRVNSCQCLAVSPMCAMSCDEALPLGSPRCDIENILKDNCAPESTEPVSEAR 62
 QY 63 VLERPLPSDKGSDSSQVPSVPSRQIARLPRDPSKPSIQVQVEDYPVVYIYMDSY 122
 Db 88 ILEARPLPSKGSSAQITQVSPQRIVRLRFDSSKFLSQVQRQVEDYPVVYIYMDSF 147
 QY 123 SMKDQDLSIIONLQGKTAKOMRKTSNLRIGAFVDPKVSPMYLSPPEAENPCYDMDKT 182
 Db 148 SMKDQDLSIOTLGKTLASQMRKLTSLRIGAFVDPKVSPMYLSPQATKNPCYDMDKT 207
 QY 183 TCAUPMFGYKHVLTLDQVTRFHEEVKSVSRNRDAPEGGDAIMQATVCDKIGRNDA 242
 Db 208 ACPLPMEGYKHVLTLDQVSRFNEVKQSVSRNRDAPEGGDAIMQATVCDKIGRNDA 267
 QY 243 SHLVLFTVDTAKTHIALDGRLAGIVPLNDGHHRGHTDHYASSTMOPSLGIMTEKLSQK 302
 Db 268 SHLVLFTVDTAKTHIALDGRLAGIVPLNDGHHRGHTDHYASSTMOPSLGIMTEKLSQK 327
 QY 303 NINLIAVTEVNVLYQYESLIPGTTVGVLSMDSSNVLOQTIVDAYKIRSKVLEVRDL 362
 Db 328 NINLIAVTEVNVLYQYESLIPGTTVGVLSMDSSNVLOQTIVDAYKIRSKVLEVRDL 387
 QY 363 PEBLSSLNATCLNEVPLGKLSMGMKIGDIVSSTEAKRGCPKEKEFTIKVGFK 422
 Db 388 PGELSLSLSPNATCLNEVPLGKLSMGMKIGDIVSSTEAKRGCPKEKEFTIKVGFK 447
 QY 423 DSELIVQVTFDCACQACQAEPAFHKRCNGNCFECGVCRCGWLGSQCESEEDYRPSQ 482
 Db 448 DSELIVQVTFDCACQACQAEPAFHKRCNGNCFECGVCRCGWLGSQCESEEDYRPSQ 507
 QY 483 QDECSSPRGQPVCSQRGBCLCQCVCHSISDRGKRTGKYCECDDFSCVRYKEMCSGHQGC 542
 Db 508 QDECSSPRGQPVCSQRGBCLCQCVCHSISDRGKRTGKYCECDDFSCVRYKEMCSGHQGC 567
 Db 529 QDECSSPRGQPVCSQRGBCLCQCVCHSISDRGKRTGKYCECDDFSCVRYKEMCSGHQGC 567
 QY 543 SGCDCLCSDSWGTYGYYCNCCTTDTCSMSNGLICSGSGKCGEGSCVCIQPSYGDPEKCP 602
 Db 568 NGCDCVCDSDWGTGYYCNCCTTDTCSMSNGLICSGRNCCEGSCVCPGSGYGDICEKCP 627
 QY 603 TEPDACTFKKCEVCKKFDGALHDENTCNCRDTEIESVELKDGDAWNCTKNEED 662

RESULT 3
 ITB5_MOUSE
 ID ITB5_MOUSE STANDARD; PRT; 798 AA.
 AC 07030; 070308; 088347; (Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Integrin beta-5 precursor.
 GN ITGB5.
 OS Mus musculus (Mouse)
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OC OX NCBI_TaxId=10090;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORMS BETA-5A AND BETA-5B).
 RC TISSUE=Liver;
 RA Zhang H., Tan S.M., Lu J.;
 RT MEDLINE=98198405; Pubmed=9531507;
 RL Zhang H., Tan S.M., Lu J.;
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORM BETA-5A).
 RC TISSUE=Brain;
 RX MEDLINE=99098874; Pubmed=9880508;
 RA Fen X., Teitelbaum S.L., Quroz M.E., Towler D.A., Ross F.P.;
 RT "Cloning of the murine beta5 integrin subunit promoter. Identification
 of a novel sequence mediating granulocyte-macrophage colony-
 stimulating factor-dependent repression of beta5 integrin gene
 transcription.",
 RT J. Biol. Chem. 274:1366-1374 (1999).
 RL J. BIOL. CHEM. 274:1366-1374 (1999).
 CC -!- FUNCTION: INTEGRIN ALPHAV/BETTA5 IS A RECEPTOR FOR FIBRONECTIN.
 CC IT RECOGNIZES THE SEQUENCE R-G-D IT ITS LIGAND.
 CC -!- SUBUNIT: HETEROPIPER OF AN ALPHAV AND A BETA SUBUNIT. BETA-5
 CC ASSOCIATES WITH ALPHAV.
 CC -!- SUBCELLULAR LOCATION: TYPE I membrane protein
 CC -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS: BETA5A (SHOWN HERE) AND BETA-
 CC 5B; ARE PRODUCED BY ALTERNATIVE SPLICING.
 CC -!- SIMILARITY: BELONGS TO THE INTEGRIN BETA CHAIN FAMILY.
 CC -!- SIMILARITY: CONTAINS 1 VWA-LIKE DOMAIN.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC
 CC
 DR EMBL; AF043257; AAC01101; -.
 DR EMBL; AF043256; AAC01091; -.
 DR EMBL; AF022110; AAD08782; 1; -.
 DR EMBL; AF03355; 2MRT.
 DR MGD; MG-1; 96614; ITGB5.
 DR InterPro; IPR000561; EGR-like.
 DR InterPro; IPR002369; Integrin_B.
 DR InterPro; IPR003659; PSI.
 DR InterPro; IPR002035; vWA.
 DR Pfam; PF00362; Integrin_B; 1.
 DR PRINTS; PR01186; INTEGRIN_B.
 DR PRODOM; PD00181; Integrin_B; 1.

RT	"Cloning of an integrin beta subunit exhibiting high homology with	FT	DISULFID	484	522	BY SIMILARITY.
RT	integrin beta 3 subunit.";	FT	DISULFID	489	498	BY SIMILARITY.
RL	"cDNA sequence of the human integrin beta 5 subunit.";	FT	DISULFID	500	513	BY SIMILARITY.
RN	[3]	FT	DISULFID	528	533	BY SIMILARITY.
RX	SEQUENCE FROM N.A.	FT	DISULFID	530	563	BY SIMILARITY.
RA	McLean J.W., Vestal D.J., Cheresh D.A., Bodary S.C.;	FT	DISULFID	535	548	BY SIMILARITY.
RT	"cDNA sequence of the human integrin beta 5 subunit.";	FT	DISULFID	550	555	BY SIMILARITY.
RL	J. Biol. Chem. 265:17126-17131 (1990).	FT	DISULFID	569	574	BY SIMILARITY.
CC	-!- FUNCTION: INTEGRIN ALPHAV/BETAV-5 IS A RECEPTOR FOR FIBRONECTIN.	FT	DISULFID	571	602	BY SIMILARITY.
CC	IT RECOGNIZES THE SEQUENCE R-G-D IT ITS LIGAND.	FT	DISULFID	576	585	BY SIMILARITY.
CC	-!- SUBUNIT: HETERO-DIMER OF AN ALPHAV AND A BETAV SUBUNIT. BETA-5	FT	DISULFID	587	594	BY SIMILARITY.
CC	ASSOCIATES WITH ALPHAV.	FT	DISULFID	608	613	BY SIMILARITY.
CC	-!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.	FT	DISULFID	610	657	BY SIMILARITY.
CC	-!- SIMILARITY: BELONGS TO THE INTEGRIN BETA CHAIN FAMILY.	FT	DISULFID	615	625	BY SIMILARITY.
CC	-!- SIMILARITY: CONTAINS 1 WVEA-LIKE DOMAIN.	FT	DISULFID	628	631	BY SIMILARITY.
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration	FT	DISULFID	635	644	BY SIMILARITY.
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -	FT	DISULFID	641	714	BY SIMILARITY.
CC	the European Bioinformatics Institute. There are no restrictions on its	FT	DISULFID	641	690	BY SIMILARITY.
CC	use by non-profit institutions as long as its content is in no way	FT	CARBOHYD	347	347	N-LINKED (GLCNAC. . .) (POTENTIAL).
CC	modified and this statement is not removed. Usage by and for commercial	FT	CARBOHYD	460	460	N-LINKED (GLCNAC. . .) (POTENTIAL).
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/) or send an email to license@isb-sib.ch).	FT	CARBOHYD	477	477	N-LINKED (GLCNAC. . .) (POTENTIAL).
CC	-----	FT	CARBOHYD	505	505	N-LINKED (GLCNAC. . .) (POTENTIAL).
CC	-----	FT	CARBOHYD	552	552	N-LINKED (GLCNAC. . .) (POTENTIAL).
CC	-----	FT	CARBOHYD	586	586	N-LINKED (GLCNAC. . .) (POTENTIAL).
DR	EMBL: X53002; CAA37188_1;	FT	CARBOHYD	654	654	N-LINKED (GLCNAC. . .) (POTENTIAL).
DR	M30011; AAA52707_1;	FT	CARBOHYD	705	705	N-LINKED (GLCNAC. . .) (POTENTIAL).
DR	EMBL: J05633; AAA59183_1;	FT	CONFLICT	193	193	T -> A (IN REF. 2).
DR	PIR: A35775; A35775.	FT	CONFLICT	645	645	L -> P (IN REF. 3).
DR	PIR: S12334; S12334.	FT	CONFLICT	790	792	MISSING (IN REF. 2).
DR	PIR: A38308; A38308.	FT	SEQUENCE	799	88053 MW; D7E4/27C3/105/2B CRC64;	
DR	PIR: S11708; S11708.					
DR	HSSP; P04355; 2MRT.					
DR	MIM: 147561; -.					
DR	Inter-Pro: IPRO00561; EGF-like.					
DR	Inter-Pro: IPRO02369; Integrin_beta_C.					
DR	Inter-Pro: IPRO0169; Integrin_beta_C.					
DR	Inter-Pro: IPRO0369; PST.					
DR	Inter-Pro: IPRO02035; WVEA.					
DR	Pfam: PF00362; integrin_B.					
DR	PRINTS: PRO1186; INTEGRINB.					
DR	PRODOM: PD00181; Integrin_B.					
DR	SMART: SM00181; EGF_1.					
DR	SMART: SM00187; INB_1.					
DR	SMART: SM00423; PST_1.					
DR	SMART: SM00327; WVA_1.					
DR	PROSITE: PS00243; INTEGRIN_BETA_2.					
DR	PROSITE: PS00022; EGF_1; UNKNOWN_2.					
KW	Integrin; Cell adhesion; Receptor; Transmembrane; Glycoprotein;					
Repeat; Signal; Repeat; Signal.	POTENTIAL.					
FT	SIGNAL_1	23	POTENTIAL.			
FT	CHAIN	24	INTEGRIN_BETA_5.			
FT	DOMAIN	24	EXTRACELLULAR (POTENTIAL).			
FT	TRANSMEM	720	POTENTIAL.			
FT	DOMAIN	743	CYTOPLASMIC (POTENTIAL).			
FT	DOMAIN	136	WVEA-LIKE.			
FT	DOMAIN	630	4 CYSTEINE-RICH TANDEM REPEATS.			
FT	REPEAT	465	I.			
FT	REPEAT	513	II.			
FT	REPEAT	555	III.			
FT	REPEAT	593	IV.			
FT	REPEAT	594	BY SIMILARITY.			
FT	DISULFID	463	BY SIMILARITY.			
FT	DISULFID	36	BY SIMILARITY.			
FT	DISULFID	39	BY SIMILARITY.			
FT	DISULFID	64	BY SIMILARITY.			
FT	DISULFID	202	BY SIMILARITY.			
FT	DISULFID	259	BY SIMILARITY.			
FT	DISULFID	401	BY SIMILARITY.			
FT	DISULFID	433	BY SIMILARITY.			
FT	DISULFID	461	BY SIMILARITY.			
FT	DISULFID	476	BY SIMILARITY.			
FT	DISULFID	487	BY SIMILARITY.			
FT	DISULFID	489	BY SIMILARITY.			
FT	DISULFID	500	BY SIMILARITY.			
FT	DISULFID	528	BY SIMILARITY.			
FT	DISULFID	530	BY SIMILARITY.			
FT	DISULFID	535	BY SIMILARITY.			
FT	DISULFID	550	BY SIMILARITY.			
FT	DISULFID	569	BY SIMILARITY.			
FT	DISULFID	571	BY SIMILARITY.			
FT	DISULFID	576	BY SIMILARITY.			
FT	DISULFID	587	BY SIMILARITY.			
FT	DISULFID	608	BY SIMILARITY.			
FT	DISULFID	610	BY SIMILARITY.			
FT	DISULFID	613	BY SIMILARITY.			
FT	DISULFID	617	BY SIMILARITY.			
FT	DISULFID	628	BY SIMILARITY.			
FT	DISULFID	635	BY SIMILARITY.			
FT	DISULFID	641	BY SIMILARITY.			
FT	DISULFID	644	BY SIMILARITY.			
FT	DISULFID	661	BY SIMILARITY.			
FT	DISULFID	674	BY SIMILARITY.			
FT	DISULFID	680	BY SIMILARITY.			
FT	DISULFID	690	BY SIMILARITY.			
FT	DISULFID	694	BY SIMILARITY.			
FT	DISULFID	714	BY SIMILARITY.			
FT	DISULFID	717	BY SIMILARITY.			
FT	DISULFID	720	BY SIMILARITY.			
FT	DISULFID	724	BY SIMILARITY.			
FT	DISULFID	728	BY SIMILARITY.			
FT	DISULFID	732	BY SIMILARITY.			
FT	DISULFID	736	BY SIMILARITY.			
FT	DISULFID	740	BY SIMILARITY.			
FT	DISULFID	744	BY SIMILARITY.			
FT	DISULFID	748	BY SIMILARITY.			
FT	DISULFID	752	BY SIMILARITY.			
FT	DISULFID	756	BY SIMILARITY.			
FT	DISULFID	760	BY SIMILARITY.			
FT	DISULFID	764	BY SIMILARITY.			
FT	DISULFID	768	BY SIMILARITY.			
FT	DISULFID	772	BY SIMILARITY.			
FT	DISULFID	776	BY SIMILARITY.			
FT	DISULFID	780	BY SIMILARITY.			
FT	DISULFID	784	BY SIMILARITY.			
FT	DISULFID	788	BY SIMILARITY.			
FT	DISULFID	792	BY SIMILARITY.			
FT	DISULFID	796	BY SIMILARITY.			
FT	DISULFID	800	BY SIMILARITY.			
FT	DISULFID	804	BY SIMILARITY.			
FT	DISULFID	808	BY SIMILARITY.			
FT	DISULFID	812	BY SIMILARITY.			
FT	DISULFID	816	BY SIMILARITY.			
FT	DISULFID	820	BY SIMILARITY.			
FT	DISULFID	824	BY SIMILARITY.			
FT	DISULFID	828	BY SIMILARITY.			
FT	DISULFID	832	BY SIMILARITY.			
FT	DISULFID	836	BY SIMILARITY.			
FT	DISULFID	840	BY SIMILARITY.			
FT	DISULFID	844	BY SIMILARITY.			
FT	DISULFID	848	BY SIMILARITY.			
FT	DISULFID	852	BY SIMILARITY.			
FT	DISULFID	856	BY SIMILARITY.			
FT	DISULFID	860	BY SIMILARITY.			
FT	DISULFID	864	BY SIMILARITY.			
FT	DISULFID	868	BY SIMILARITY.			
FT	DISULFID	872	BY SIMILARITY.			
FT	DISULFID	876	BY SIMILARITY.			
FT	DISULFID	880	BY SIMILARITY.			
FT	DISULFID	884	BY SIMILARITY.			
FT	DISULFID	888	BY SIMILARITY.			
FT	DISULFID	892	BY SIMILARITY.			
FT	DISULFID	896	BY SIMILARITY.			
FT	DISULFID	900	BY SIMILARITY.			
FT	DISULFID	904	BY SIMILARITY.			
FT	DISULFID	908	BY SIMILARITY.			
FT	DISULFID	912	BY SIMILARITY.			
FT	DISULFID	916	BY SIMILARITY.			
FT	DISULFID	920	BY SIMILARITY.			
FT	DISULFID	924	BY SIMILARITY.			
FT	DISULFID	928	BY SIMILARITY.			
FT	DISULFID	932	BY SIMILARITY.			
FT	DISULFID	936	BY SIMILARITY.			
FT	DISULFID	940	BY SIMILARITY.			
FT	DISULFID	944	BY SIMILARITY.			
FT	DISULFID	948	BY SIMILARITY.			
FT	DISULFID	952	BY SIMILARITY.			
FT	DISULFID	956	BY SIMILARITY.			
FT	DISULFID	960	BY SIMILARITY.			
FT	DISULFID	964	BY SIMILARITY.			
FT	DISULFID	968	BY SIMILARITY.			
FT	DISULFID	972	BY SIMILARITY.			
FT	DISULFID	976	BY SIMILARITY.			
FT	DISULFID	980	BY SIMILARITY.			
FT	DISULFID	984	BY SIMILARITY.			
FT	DISULFID	988	BY SIMILARITY.			
FT	DISULFID	992	BY SIMILARITY.			
FT	DISULFID	996	BY SIMILARITY.			
FT	DISULFID	1000	BY SIMILARITY.			
FT	DISULFID	1004	BY SIMILARITY.			
FT	DISULFID	1008	BY SIMILARITY.			
FT	DISULFID	1012	BY SIMILARITY.			
FT	DISULFID	1016	BY SIMILARITY.			
FT	DISULFID	1020	BY SIMILARITY.			
FT	DISULFID	1024	BY SIMILARITY.			
FT	DISULFID	1028	BY SIMILARITY.			
FT	DISULFID	1032	BY SIMILARITY.			
FT	DISULFID	1036	BY SIMILARITY.			
FT	DISULFID	1040	BY SIMILARITY.			
FT	DISULFID	1044	BY SIMILARITY.			
FT	DISULFID	1048	BY SIMILARITY.			
FT	DISULFID	1052	BY SIMILARITY.			
FT	DISULFID	1056	BY SIMILARITY.			
FT	DISULFID	1060	BY SIMILARITY.			
FT	DISULFID	1064	BY SIMILARITY.			
FT	DISULFID	1068	BY SIMILARITY.			
FT	DISULFID	1072	BY SIMILARITY.			
FT	DISULFID	1076	BY SIMILARITY.			
FT	DISULFID	1080	BY SIMILARITY.			
FT	DISULFID	1084	BY SIMILARITY.			
FT	DISULFID	1088	BY SIMILARITY.			
FT	DISULFID	1092	BY SIMILARITY.			
FT	DISULFID	1096	BY SIMILARITY.			
FT	DISULFID	1100	BY SIMILARITY.			
FT	DISULFID	1104	BY SIMILARITY.			
FT	DISULFID	1108	BY SIMILARITY.			
FT	DISULFID	1112	BY SIMILARITY.			
FT	DISULFID	1116	BY SIMILARITY.			
FT	DISULFID	1120	BY SIMILARITY.			
FT	DISULFID	1124	BY SIMILARITY.			
FT	DISULFID	1128	BY SIMILARITY.			
FT	DISULFID	1132	BY SIMILARITY.			
FT	DISULFID	1136	BY SIMILARITY.			
FT	DISULFID	1140	BY SIMILARITY.			
FT	DISULFID	1144	BY SIMILARITY.			
FT	DISULFID	1148	BY SIMILARITY.			
FT	DISULFID	1152	BY SIMILARITY.			
FT	DISULFID	1156	BY SIMILARITY.			
FT	DISULFID	1160	BY SIMILARITY.			
FT	DISULFID	1164				

RESULT 5

ID ITGB_MOUSE STANDARD PRY: 787 AA.

ITGB_MOUSE SIGNAL 1

AC FT SIGNAL 21

OZG079; FT CHAIN 21

16-OCT-2001 (Rel. 40, Created) FT DOMAIN 22

DT 16-OCT-2001 (Rel. 40, Last sequence update) FT TRANSMEM 706

01-MAR-2002 (Rel. 41, Last annotation update) FT DOMAIN 707

DE Integrin beta-6 precursor, FT DOMAIN 787

ITGB6. FT DOMAIN 771

OS Mus musculus (Mouse). FT DOMAIN 456

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; FT DISULFD 619

OX Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. FT DISULFD 501

{1} FT DISULFD 543

RP TISSUE=Kidney; FT DISULFD 582

RX MEDLINE=20547423; PubMed=11095652; FT DISULFD 583

RA Arond J., Smart A.M., Briggs J.P.; FT DISULFD 594

RT "Mouse beta(6) integrin sequence, pattern of expression, and role in FT DISULFD 595

Kidney development"; FT DISULFD 596

RL Soc Nephrol, 11:2297-2305(2000). FT DISULFD 597

CC - - FUNCTION: INTEGRIN ALPHAI-BETA-6 IS A RECEPTOR FOR FIBRONECTIN FT DISULFD 598

AND CYTOACTIN. IT RECOGNIZES THE SEQUENCE R-G-D IN ITS LIGANDS FT DISULFD 599

CC (BY SIMILARITY). FT DISULFD 600

CC - - SUBUNIT: HETERO-DIMER OF AN ALPHA AND A BETA SUBUNIT. BETA-6 FT DISULFD 601

ASSOCIATES WITH ALPHA-V (BY SIMILARITY). FT DISULFD 602

CC - - SUBCELLULAR LOCATION: TYPE I membrane protein. FT DISULFD 603

CC - - SIMILARITY: BELONGS TO THE INTEGRIN BETA CHAIN FAMILY. FT DISULFD 604

CC - - SIMILARITY: CONTAINS 1 WFPA-LIKE DOMAIN. FT DISULFD 605

This SWISS-PROT entry is copyright between the Swiss Institute of Bioinformatics and the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or_send_an_email_to_license@isb-sib.ch).

DR MGD: MG1:96635; ITGB6. FT DISULFD 606

DR InterPro; IPR00561; EGF-like. FT DISULFD 607

DR InterPro; IPR002369; Integrin_B. FT DISULFD 608

DR InterPro; IPR001169; Integrin_beta_C. FT DISULFD 609

DR InterPro; IPR003559; PSI. FT DISULFD 610

DR InterPro; IPR002035; WFPA. FT DISULFD 611

DR Pfam; PF00327; Integrin_B. FT DISULFD 612

DR PRINTS; P01186; INTEGRINB. FT DISULFD 613

DR SMART; SM00021; EGF-like. FT DISULFD 614

DR SMART; SM00187; INB. FT DISULFD 615

DR SMART; SM00423; PSI. FT DISULFD 616

DR SMART; SM00327; WFPA. FT DISULFD 617

DR PRINTS; P00181; INTEGRIN_B. FT DISULFD 618

DR SMART; SM00021; EGF-like. FT DISULFD 619

DR SMART; SM00187; INB. FT DISULFD 620

DR SMART; SM00423; PSI. FT DISULFD 621

DR SMART; SM00327; WFPA. FT DISULFD 622

DR PRINTS; P00184; INTEGRIN_B. FT DISULFD 623

DR SMART; SM00022; EGF-like. FT DISULFD 624

DR SMART; PS01186; EGF_2. UNKNOWN_1. FT DISULFD 625

Integrin; Cell adhesion; Receptor; Transmembrane; Glycoprotein; Repeat; Signal.

EMBL: AB115376; ADP1712.1. FT DISULFD 626

Query Match Best Local Similarity 52.4%; Score 2177; DB 1; Length 787; Matches 392; Conservative 51.6%; pred. No. 4.2e-140; Mismatches 232; Inels 12; Gaps 7;

Matches 392; Conservative 51.6%; pred. No. 4.2e-140; Mismatches 232; Inels 12; Gaps 7;

QY 5 CTTGGSSCOOCLAVSPMWCACSD--ALPLGSPRCDLKEUNIJKDNFAPESLEFPYSEA 61

Db 23 CAWGAESCDCLLGPAGACSGENFHLSGAGERCOTPANLAKGQQLFTENPVSRI 82

QY 62 RYLDERPLSKSGSDSSQVQVSPIAQLRLRPPDKSKFNSIVQROVEYPDVYMLDS 121

Db 83 EVLQNKPLSLVGQRKNSSDVIQIAPLSVLKRPGREQTIVQVQRTEDYPVDLVYMLDS 142

QY 122 YSMKDLWSTQNGKPLAQYQMKLTSNLRIGFGAFVDPVSPWYISPEALENPYCDMK 181

Db 143 ASMDDDLNITKELGSLRAKMSKLTNSRFLGFSFVEKPSVPMK-TPEEELTPNCSSIP 201

QY 182 TTCLPMFGKVKHLTLDQTRFNEEVKKQOSVRNDAPEEGFIMQKVCDEKIGRND 241

Db 202 YFCLOTFEGHFLPLTDFERENLYRQKQKISANIDTREGFAPIMQAVCKEKGIRD 261

QY 242 ASHLVFTTDAAKTHALDGRLAGLIVOPNDGQCHVGSNDHYSASITMDYPSLGMTELSQ 301

Db 262 SLHLVVFSDADSHFGMDSKLAGIVIPNDGICLHDHRNEYSMSTVLEYPITQIQLIDKLVQ 321
 CC - ! - SUBCELLULAR LOCATION: TYPE I membrane protein.
 QY 302 KNINLIFAVTENVNLYQNTISELIPTTVGVLMSOSSNVQOLIVDAYGKTRSKVELERD 361
 CC - ! - SIMILARITY: BELONGS TO THE INTEGRIN BETA CHAIN FAMILY.
 Db 322 NNVLLIFAVTQEQQVHLYENYAKLICATGVGLQKSGNQNLQILISAYEEERSEVELEUG 381
 CC - ! - SIMILARITY: CONTAINS 1 WWF-A-LIKE DOMAIN.
 QY 362 LPEELSLSFNATCLNNENVTPQIKSCMGKIKDTVSFSIEAKVRGCPQEKEKFTKPVGF 421
 CC - ! - SIMILARITY: This SWISS-PROT entry is copyright. It is produced through a collaboration
 Db 382 DTEGLNLNSFTALCNGVLFQPKKCSHMKVGDIAFSNVTWSVSVS-EKRSRNLKPKVGL 440
 CC - ! - SIMILARITY: between the Swiss Institute of Bioinformatics and the EMBL outstation
 QY 422 KSLIVYQVTFDCDCAQAOAQPAPNSHNCNGGTERCGVCGPGWGLSGOCCESEDYRPS 481
 CC - ! - SIMILARITY: the European Bioinformatics Institute. There are no restrictions on its
 Db 441 GDTELELIVSAECDCDCQREIETPNSSKCHNGHSFOQGVCTNPGHMPHCECGED--MV 497
 CC - ! - SIMILARITY: use by non-profit institutions as long as its content is in no way
 QY 482 QODECSPREGPVCPSORGECQGCVCHSSDFGKIGTKYCECDDFSCVRYGEMSGHQQ 541
 CC - ! - SIMILARITY: modified and this statement is not removed. Usage by and for commercial
 Db 498 SADSCKSRESPGPGPSGCGDVCQGOCPICHLPSYGSITGPGCQCDNFSCLRHKGLLGDNGD 557
 CC - ! - SIMILARITY: entities requires a license agreement (see <http://www.isb-sib.ch/announce/>)
 QY 542 CSGGDCCLCDSDPWTGYYCNCCTRDTOMSSNGLLCSRGKCCGSCVCIOPGSYGYDCEK 601
 CC - ! - SIMILARITY: or send an email to license@ib-sib.ch).
 Db 558 CPGCFCVCRDWTGECNCNTNRDSCTSEDQVLCGSGDVCVKGKCVRNPGGSGETCERC 617
 CC - ! - SIMILARITY:
 QY 602 PPGCPDRAFTKFCVCEKVECKFDIGALHIDENTCVCYRCDEIEVSKEELKDGTGK-AVNCYTKNE 660
 CC - ! - SIMILARITY:
 Db 618 PTGCDPENPSKRCSCIECKLSDAQAOEE--CDKDKRAIGATI-SE-EDFSKDISVSLOGE 674
 CC - ! - SIMILARITY:
 QY 661 DCVWRQYQYEDSGSKSILVYVEPPCPKGDIYLVLSSYGAILLGLAALLWKLIT 720
 CC - ! - SIMILARITY:
 Db 675 NRECLITFLITIDNEKGTHIHNNEKDCPKPPNIPMLGVSLAILLIGWVLLCTWKLVS 734
 CC - ! - SIMILARITY:
 QY 721 THDRKEVAKFEERARAKWDTANPPLYKEATSTFTNITYR 760
 CC - ! - SIMILARITY:
 Db 735 FHDKEVAKFEERSKAKWQGTNPPLYRGSTSTEKNVYK 774
 CC - ! - SIMILARITY:
 RESULT 6
 ITB6_HUMAN STANDARD; PRT; 788 AA.
 AC P18564; Q16500; 01-NOV-1990 (Rel. 16, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Integrin beta-6 precursor.
 GN ITGB6.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID:9606;
 [1] SEQUENCE FROM N.A.
 RP TISSUE-Pancreas;
 RX MEDLINE=90307659; PubMed=2365683;
 RA Sheppard D., Rozzo C., Starr L., Quaranta V., Erle D.J., Pytela R.;
 RT Complete amino acid sequence of a novel integrin beta subunit (beta
 6) identified in epithelial cells using the polymerase chain
 reaction., J. Biol. Chem. 265:11502-11507(1990).
 [2] REVISONS TO 18-24; 158; 642 AND 719.
 RA Askins J.;
 RL Submitted (SEP-2000) to the EMBL/GenBank/DDBJ databases.
 [3] SEQUENCE OF 116-197 FROM N.A.
 RX MEDLINE=93002753; PubMed=1382574;
 RA Jiang W.M., Jenkins D., Yuan Q., Leung E., Choo K.H., Watson J.D.,
 RA Krissansen G.W.;
 "The gene organization of the human beta 7 subunit, the common beta
 subunit of the leukocyte integrins HML-1 and LPM-1.";
 RL Int. Immunol. 4:1031-1040(1992).
 CC - ! - FUNCTION: INTEGRIN ALPHAI- β 7-6 IS A RECEPTOR FOR FIBRONECTIN
 CC AND CYTOKINETIC ITT RECOGNIZES THE SEQUENCE R-G-D IT ITS LIGANDS.
 CC - ! - SUBUNIT: HETEROODIMER OF AN ALPHAI AND A BETA SUBUNIT. BETA-6
 CC ASSOCIATES WITH ALPHAI-7.
 ASSOCIATES WITH ALPHAI-7.

CC - ! - SUBCELLULAR LOCATION: TYPE I membrane protein.
 CC - ! - SIMILARITY: BELONGS TO THE INTEGRIN BETA CHAIN FAMILY.
 CC - ! - SIMILARITY: CONTAINS 1 WWF-A-LIKE DOMAIN.
 This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>)
 or send an email to license@ib-sib.ch).
 DR EMBL; M35198; AAA36122.; --
 DR EMBL; A3609; CAA01832.; --
 DR EMBL; S49380; AAB36901.; --
 DR PIR; A37057; A37057.
 DR HSSP; P03355; 2MRT.
 DR MM: 14758; --
 DR InterPro; IPR00561; EGFR-LIKE.
 DR InterPro; IPR02369; Integrin_B.
 DR InterPro; IPR03659; PSI.
 DR InterPro; IPR002035; VWA.
 DR Pfam; PF00362; Integrin_B; 1.
 DR Pfam; PR01186; INTEGRIN_B.
 DR Prodom; PD001811; Integrin_B; 1.
 DR SMART; SM00001; EGF-LIKE; 1.
 DR SMART; SM00187; INB; 1.
 DR SMART; SM00423; ISI; 1.
 DR SMART; SM00327; VWA; 1.
 DR PROSITE; PS000243; INTEGRIN BETA; 3.
 DR PROSITE; PS01186; EGF2; UNKNOWN_1.
 DR PROSITE; PS01186; EGFR; UNKNOWN_1.
 KW Integrin; Cell adhesion; Receptor; Transmembrane; Glycoprotein;
 KW Repeat; Signal; 1
 FT SIGNAL 1 21 POTENTIAL.
 FT SIGNAL 22 788 INTEGRIN BETA-6.
 FT DOMAIN 22 707 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 708 730 POTENTIAL.
 FT DOMAIN 731 788 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 131 371 VWA-LIKE.
 FT DOMAIN 456 619 4 CSTEINE-RICH TANDEM REPEATS.
 FT REPEAT 456 501 I.
 FT REPEAT 502 543 II.
 FT REPEAT 544 582 III.
 FT REPEAT 583 619 IV.
 FT DISULFID 23 454 BY SIMILARITY.
 FT DISULFID 31 41 BY SIMILARITY.
 FT DISULFID 34 70 BY SIMILARITY.
 FT DISULFID 34 70 BY SIMILARITY.
 FT DISULFID 44 59 BY SIMILARITY.
 FT DISULFID 197 204 BY SIMILARITY.
 FT DISULFID 252 293 BY SIMILARITY.
 FT DISULFID 394 406 BY SIMILARITY.
 FT DISULFID 426 670 BY SIMILARITY.
 FT DISULFID 452 456 BY SIMILARITY.
 FT DISULFID 467 479 BY SIMILARITY.
 FT DISULFID 476 511 BY SIMILARITY.
 FT DISULFID 481 490 BY SIMILARITY.
 FT DISULFID 492 502 BY SIMILARITY.
 FT DISULFID 517 522 BY SIMILARITY.
 FT DISULFID 519 552 BY SIMILARITY.
 FT DISULFID 524 537 BY SIMILARITY.
 FT DISULFID 539 544 BY SIMILARITY.
 FT DISULFID 558 563 BY SIMILARITY.
 FT DISULFID 560 591 BY SIMILARITY.
 FT DISULFID 565 574 BY SIMILARITY.
 FT DISULFID 576 583 BY SIMILARITY.
 FT DISULFID 597 602 BY SIMILARITY.
 FT DISULFID 645 645 BY SIMILARITY.
 FT DISULFID 659 664 BY SIMILARITY.
 FT DISULFID 664 674 BY SIMILARITY.
 FT DISULFID 677 677 BY SIMILARITY.
 FT DISULFID 677 680 BY SIMILARITY.
 FT DISULFID 684 693 BY SIMILARITY.
 FT DISULFID 693 702 BY SIMILARITY.
 FT DISULFID 693 702 BY SIMILARITY.

Query Match 52.3%; Score 2170.5; DB 1; Length 788; Best Local Similarity 50.9%; Pref. No. 1-2a-139; Matches 387; Conservative 134; Mismatches 228; Indels 11; Gaps 7; N-LINKED (GLCNAC. . .) (POTENTIAL). FT CARBOHYD 48 48 FT CARBOHYD 97 97 FT CARBOHYD 260 260 FT CARBOHYD 387 387 FT CARBOHYD 396 396 FT CARBOHYD 463 463 FT CARBOHYD 471 471 FT CARBOHYD 541 541 FT CARBOHYD 575 575 FT CARBOHYD 788 AA; 85935 MW; ED7D7533EC4C8C4D CRC64; SQ

Query Match 52.3%; Score 2170.5; DB 1; Length 788; Best Local Similarity 50.9%; Pref. No. 1-2a-139; Matches 387; Conservative 134; Mismatches 228; Indels 11; Gaps 7; N-LINKED (GLCNAC. . .) (POTENTIAL). FT CARBOHYD 97 97 FT CARBOHYD 260 260 FT CARBOHYD 387 387 FT CARBOHYD 396 396 FT CARBOHYD 463 463 FT CARBOHYD 471 471 FT CARBOHYD 541 541 FT CARBOHYD 575 575 FT CARBOHYD 788 AA; 85935 MW; ED7D7533EC4C8C4D CRC64; SQ

RP SEQUENCE FROM N-A. RX MEDLINE=9440831; PubMed=8224922; RA Shoji M., Hayzer D.J., Kim T.M., Runge M.S., Hanson S.R.; RT "Human and baboon integrin beta 5 subunit-encoding mRNAs have RT alternative polyadenylation sites.", RL Gene 133:307-308(1993). CC -I- FUNCTION: INTEGRIN ALPHAV-BETA5 IS A RECEPTOR FOR FIBRONECTIN CC -I- SUBUNIT: HETEROODIMER OF AN ALPHAV AND A BETA SUBUNIT. BETA5 CC -I- SUBCELLULAR LOCATION: TYPE I membrane protein. CC -I- PIM: THE CYSTEINE RESIDUES ARE INVOLVED IN INTRACHAIN DISULFIDE BONDS. CC -I- SIMILARITY: BELONGS TO THE INTEGRIN BETA CHAIN FAMILY. CC -I- SIMILARITY: CONTAINS 1 WW-like domain.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation, the European Bioinformatics Institute. There are no restrictions on use by non-profit institutions as long as its content is in modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/ansatz> or send an email to license@isb-sib.ch).

CC DR EMBL: L12231; AAA16866.1; -. CC DR PRODOM: P04355; 2NRT. CC DR SMART: SM00561; EGF-like. CC DR SMART: SM00187; INB; 1. CC DR SMART: SM00327; WWA; 1. CC DR Interpro: IPR003369; Integrin_B. CC DR Interpro: IPR001169; Integrin_beta_C. CC DR Pfam: PF00362; Integrin_B; 1. CC DR PRODOM: P001811; Integrin_B; 1. CC DR EMBL: L12231; AAA16866.1; -. CC DR PRODOM: P001811; Integrin_B; 1. CC DR SMART: SM00561; EGF-like. CC DR SMART: SM00187; INB; 1. CC DR SMART: SM00327; WWA; 1. CC DR PROSITE: PS00243; INTEGRIN_BETA; 2. CC DR PROSITE: PS00022; EGF_1; UNKNOWN_2. CC DR PROSITE: PS0116; EGF_2; UNKNOWN_2. CC DR Integrin; Cell adhesion; Receptor; Transmembrane; Glycoprotein; KW Repeat; Extracellular matrix; Cytoskeleton. CC DR NON_TER 1 1 FT DOMAIN <1 575 FT DOMAIN 576 598 FT TRANSMEM 599 655 FT DOMAIN <1 234 FT DOMAIN 457 621 FT DOMAIN 321 368 FT REPEAT 369 410 FT REPEAT 411 449 FT REPEAT 450 486 FT REPEAT 451 486 FT CARBOHYD 203 203 FT CARBOHYD 316 316 FT CARBOHYD 408 408 FT CARBOHYD 442 442 FT CARBOHYD 510 510 FT CARBOHYD 561 561 FT CARBOHYD 575 575 FT CARBOHYD 599 599 FT CARBOHYD 617 617 FT CARBOHYD 621 621 FT CARBOHYD 636 636 FT CARBOHYD 640 640 FT CARBOHYD 642 642 FT CARBOHYD 650 650 FT CARBOHYD 655 655 FT CARBOHYD 660 660 FT CARBOHYD 661 661 FT CARBOHYD 662 662 FT CARBOHYD 663 663 FT CARBOHYD 664 664 FT CARBOHYD 665 665 FT CARBOHYD 666 666 FT CARBOHYD 667 667 FT CARBOHYD 668 668 FT CARBOHYD 669 669 FT CARBOHYD 670 670 FT CARBOHYD 671 671 FT CARBOHYD 672 672 FT CARBOHYD 673 673 FT CARBOHYD 674 674 FT CARBOHYD 675 675 FT CARBOHYD 676 676 FT CARBOHYD 677 677 FT CARBOHYD 678 678 FT CARBOHYD 679 679 FT CARBOHYD 680 680 FT CARBOHYD 681 681 FT CARBOHYD 682 682 FT CARBOHYD 683 683 FT CARBOHYD 684 684 FT CARBOHYD 685 685 FT CARBOHYD 686 686 FT CARBOHYD 687 687 FT CARBOHYD 688 688 FT CARBOHYD 689 689 FT CARBOHYD 690 690 FT CARBOHYD 691 691 FT CARBOHYD 692 692 FT CARBOHYD 693 693 FT CARBOHYD 694 694 FT CARBOHYD 695 695 FT CARBOHYD 696 696 FT CARBOHYD 697 697 FT CARBOHYD 698 698 FT CARBOHYD 699 699 FT CARBOHYD 700 700 FT CARBOHYD 701 701 FT CARBOHYD 702 702 FT CARBOHYD 703 703 FT CARBOHYD 704 704 FT CARBOHYD 705 705 FT CARBOHYD 706 706 FT CARBOHYD 707 707 FT CARBOHYD 708 708 FT CARBOHYD 709 709 FT CARBOHYD 710 710 FT CARBOHYD 711 711 FT CARBOHYD 712 712 FT CARBOHYD 713 713 FT CARBOHYD 714 714 FT CARBOHYD 715 715 FT CARBOHYD 716 716 FT CARBOHYD 717 717 FT CARBOHYD 718 718 FT CARBOHYD 719 719 FT CARBOHYD 720 720 FT CARBOHYD 721 721 FT CARBOHYD 722 722 FT CARBOHYD 723 723 FT CARBOHYD 724 724 FT CARBOHYD 725 725 FT CARBOHYD 726 726 FT CARBOHYD 727 727 FT CARBOHYD 728 728 FT CARBOHYD 729 729 FT CARBOHYD 730 730 FT CARBOHYD 731 731 FT CARBOHYD 732 732 FT CARBOHYD 733 733 FT CARBOHYD 734 734 FT CARBOHYD 735 735 FT CARBOHYD 736 736 FT CARBOHYD 737 737 FT CARBOHYD 738 738 FT CARBOHYD 739 739 FT CARBOHYD 740 740 FT CARBOHYD 741 741 FT CARBOHYD 742 742 FT CARBOHYD 743 743 FT CARBOHYD 744 744 FT CARBOHYD 745 745 FT CARBOHYD 746 746 FT CARBOHYD 747 747 FT CARBOHYD 748 748 FT CARBOHYD 749 749 FT CARBOHYD 750 750 FT CARBOHYD 751 751 FT CARBOHYD 752 752 FT CARBOHYD 753 753 FT CARBOHYD 754 754 FT CARBOHYD 755 755 FT CARBOHYD 756 756 FT CARBOHYD 757 757 FT CARBOHYD 758 758 FT CARBOHYD 759 759 FT CARBOHYD 760 760 FT CARBOHYD 761 761 FT CARBOHYD 762 762 FT CARBOHYD 763 763 FT CARBOHYD 764 764 FT CARBOHYD 765 765 FT CARBOHYD 766 766 FT CARBOHYD 767 767 FT CARBOHYD 768 768 FT CARBOHYD 769 769 FT CARBOHYD 770 770 FT CARBOHYD 771 771 FT CARBOHYD 772 772 FT CARBOHYD 773 773 FT CARBOHYD 774 774 FT CARBOHYD 775 775 FT CARBOHYD 776 776 FT CARBOHYD 777 777 FT CARBOHYD 778 778 FT CARBOHYD 779 779 FT CARBOHYD 780 780 FT CARBOHYD 781 781 FT CARBOHYD 782 782 FT CARBOHYD 783 783 FT CARBOHYD 784 784 FT CARBOHYD 785 785 FT CARBOHYD 786 786 FT CARBOHYD 787 787 FT CARBOHYD 788 788 FT CARBOHYD 789 789 FT CARBOHYD 790 790 FT CARBOHYD 791 791 FT CARBOHYD 792 792 FT CARBOHYD 793 793 FT CARBOHYD 794 794 FT CARBOHYD 795 795 FT CARBOHYD 796 796 FT CARBOHYD 797 797 FT CARBOHYD 798 798 FT CARBOHYD 799 799 FT CARBOHYD 800 800 FT CARBOHYD 801 801 FT CARBOHYD 802 802 FT CARBOHYD 803 803 FT CARBOHYD 804 804 FT CARBOHYD 805 805 FT CARBOHYD 806 806 FT CARBOHYD 807 807 FT CARBOHYD 808 808 FT CARBOHYD 809 809 FT CARBOHYD 810 810 FT CARBOHYD 811 811 FT CARBOHYD 812 812 FT CARBOHYD 813 813 FT CARBOHYD 814 814 FT CARBOHYD 815 815 FT CARBOHYD 816 816 FT CARBOHYD 817 817 FT CARBOHYD 818 818 FT CARBOHYD 819 819 FT CARBOHYD 820 820 FT CARBOHYD 821 821 FT CARBOHYD 822 822 FT CARBOHYD 823 823 FT CARBOHYD 824 824 FT CARBOHYD 825 825 FT CARBOHYD 826 826 FT CARBOHYD 827 827 FT CARBOHYD 828 828 FT CARBOHYD 829 829 FT CARBOHYD 830 830 FT CARBOHYD 831 831 FT CARBOHYD 832 832 FT CARBOHYD 833 833 FT CARBOHYD 834 834 FT CARBOHYD 835 835 FT CARBOHYD 836 836 FT CARBOHYD 837 837 FT CARBOHYD 838 838 FT CARBOHYD 839 839 FT CARBOHYD 840 840 FT CARBOHYD 841 841 FT CARBOHYD 842 842 FT CARBOHYD 843 843 FT CARBOHYD 844 844 FT CARBOHYD 845 845 FT CARBOHYD 846 846 FT CARBOHYD 847 847 FT CARBOHYD 848 848 FT CARBOHYD 849 849 FT CARBOHYD 850 850 FT CARBOHYD 851 851 FT CARBOHYD 852 852 FT CARBOHYD 853 853 FT CARBOHYD 854 854 FT CARBOHYD 855 855 FT CARBOHYD 856 856 FT CARBOHYD 857 857 FT CARBOHYD 858 858 FT CARBOHYD 859 859 FT CARBOHYD 860 860 FT CARBOHYD 861 861 FT CARBOHYD 862 862 FT CARBOHYD 863 863 FT CARBOHYD 864 864 FT CARBOHYD 865 865 FT CARBOHYD 866 866 FT CARBOHYD 867 867 FT CARBOHYD 868 868 FT CARBOHYD 869 869 FT CARBOHYD 870 870 FT CARBOHYD 871 871 FT CARBOHYD 872 872 FT CARBOHYD 873 873 FT CARBOHYD 874 874 FT CARBOHYD 875 875 FT CARBOHYD 876 876 FT CARBOHYD 877 877 FT CARBOHYD 878 878 FT CARBOHYD 879 879 FT CARBOHYD 880 880 FT CARBOHYD 881 881 FT CARBOHYD 882 882 FT CARBOHYD 883 883 FT CARBOHYD 884 884 FT CARBOHYD 885 885 FT CARBOHYD 886 886 FT CARBOHYD 887 887 FT CARBOHYD 888 888 FT CARBOHYD 889 889 FT CARBOHYD 890 890 FT CARBOHYD 891 891 FT CARBOHYD 892 892 FT CARBOHYD 893 893 FT CARBOHYD 894 894 FT CARBOHYD 895 895 FT CARBOHYD 896 896 FT CARBOHYD 897 897 FT CARBOHYD 898 898 FT CARBOHYD 899 899 FT CARBOHYD 900 900 FT CARBOHYD 901 901 FT CARBOHYD 902 902 FT CARBOHYD 903 903 FT CARBOHYD 904 904 FT CARBOHYD 905 905 FT CARBOHYD 906 906 FT CARBOHYD 907 907 FT CARBOHYD 908 908 FT CARBOHYD 909 909 FT CARBOHYD 910 910 FT CARBOHYD 911 911 FT CARBOHYD 912 912 FT CARBOHYD 913 913 FT CARBOHYD 914 914 FT CARBOHYD 915 915 FT CARBOHYD 916 916 FT CARBOHYD 917 917 FT CARBOHYD 918 918 FT CARBOHYD 919 919 FT CARBOHYD 920 920 FT CARBOHYD 921 921 FT CARBOHYD 922 922 FT CARBOHYD 923 923 FT CARBOHYD 924 924 FT CARBOHYD 925 925 FT CARBOHYD 926 926 FT CARBOHYD 927 927 FT CARBOHYD 928 928 FT CARBOHYD 929 929 FT CARBOHYD 930 930 FT CARBOHYD 931 931 FT CARBOHYD 932 932 FT CARBOHYD 933 933 FT CARBOHYD 934 934 FT CARBOHYD 935 935 FT CARBOHYD 936 936 FT CARBOHYD 937 937 FT CARBOHYD 938 938 FT CARBOHYD 939 939 FT CARBOHYD 940 940 FT CARBOHYD 941 941 FT CARBOHYD 942 942 FT CARBOHYD 943 943 FT CARBOHYD 944 944 FT CARBOHYD 945 945 FT CARBOHYD 946 946 FT CARBOHYD 947 947 FT CARBOHYD 948 948 FT CARBOHYD 949 949 FT CARBOHYD 950 950 FT CARBOHYD 951 951 FT CARBOHYD 952 952 FT CARBOHYD 953 953 FT CARBOHYD 954 954 FT CARBOHYD 955 955 FT CARBOHYD 956 956 FT CARBOHYD 957 957 FT CARBOHYD 958 958 FT CARBOHYD 959 959 FT CARBOHYD 960 960 FT CARBOHYD 961 961 FT CARBOHYD 962 962 FT CARBOHYD 963 963 FT CARBOHYD 964 964 FT CARBOHYD 965 965 FT CARBOHYD 966 966 FT CARBOHYD 967 967 FT CARBOHYD 968 968 FT CARBOHYD 969 969 FT CARBOHYD 970 970 FT CARBOHYD 971 971 FT CARBOHYD 972 972 FT CARBOHYD 973 973 FT CARBOHYD 974 974 FT CARBOHYD 975 975 FT CARBOHYD 976 976 FT CARBOHYD 977 977 FT CARBOHYD 978 978 FT CARBOHYD 979 979 FT CARBOHYD 980 980 FT CARBOHYD 981 981 FT CARBOHYD 982 982 FT CARBOHYD 983 983 FT CARBOHYD 984 984 FT CARBOHYD 985 985 FT CARBOHYD 986 986 FT CARBOHYD 987 987 FT CARBOHYD 988 988 FT CARBOHYD 989 989 FT CARBOHYD 990 990 FT CARBOHYD 991 991 FT CARBOHYD 992 992 FT CARBOHYD 993 993 FT CARBOHYD 994 994 FT CARBOHYD 995 995 FT CARBOHYD 996 996 FT CARBOHYD 997 997 FT CARBOHYD 998 998 FT CARBOHYD 999 999 FT CARBOHYD 1000 1000 FT CARBOHYD 1001 1001 FT CARBOHYD 1002 1002 FT CARBOHYD 1003 1003 FT CARBOHYD 1004 1004 FT CARBOHYD 1005 1005 FT CARBOHYD 1006 1006 FT CARBOHYD 1007 1007 FT CARBOHYD 1008 1008 FT CARBOHYD 1009 1009 FT CARBOHYD 1010 1010 FT CARBOHYD 1011 1011 FT CARBOHYD 1012 1012 FT CARBOHYD 1013 1013 FT CARBOHYD 1014 1014 FT CARBOHYD 1015 1015 FT CARBOHYD 1016 1016 FT CARBOHYD 1017 1017 FT CARBOHYD 1018 1018 FT CARBOHYD 1019 1019 FT CARBOHYD 1020 1020 FT CARBOHYD 1021 1021 FT CARBOHYD 1022 1022 FT CARBOHYD 1023 1023 FT CARBOHYD 1024 1024 FT CARBOHYD 1025 1025 FT CARBOHYD 1026 1026 FT CARBOHYD 1027 1027 FT CARBOHYD 1028 1028 FT CARBOHYD 1029 1029 FT CARBOHYD 1030 1030 FT CARBOHYD 1031 1031 FT CARBOHYD 1032 1032 FT CARBOHYD 1033 1033 FT CARBOHYD 1034 1034 FT CARBOHYD 1035 1035 FT CARBOHYD 1036 1036 FT CARBOHYD 1037 1037 FT CARBOHYD 1038 1038 FT CARBOHYD 1039 1039 FT CARBOHYD 1040 1040 FT CARBOHYD 1041 1041 FT CARBOHYD 1042 1042 FT CARBOHYD 1043 1043 FT CARBOHYD 1044 1044 FT CARBOHYD 1045 1045 FT CARBOHYD 1046 1046 FT CARBOHYD 1047 1047 FT CARBOHYD 1048 1048 FT CARBOHYD 1049 1049 FT CARBOHYD 1050 1050 FT CARBOHYD 1051 1051 FT CARBOHYD 1052 1052 FT CARBOHYD 1053 1053 FT CARBOHYD 1054 1054 FT CARBOHYD 1055 1055 FT CARBOHYD 1056 1056 FT CARBOHYD 1057 1057 FT CARBOHYD 1058 1058 FT CARBOHYD 1059 1059 FT CARBOHYD 1060 1060 FT CARBOHYD 1061 1061 FT CARBOHYD 1062 1062 FT CARBOHYD 1063 1063 FT CARBOHYD 1064 1064 FT CARBOHYD 1065 1065 FT CARBOHYD 1066 1066 FT CARBOHYD 1067 1067 FT CARBOHYD 1068 1068 FT CARBOHYD 1069 1069 FT CARBOHYD 1070 1070 FT CARBOHYD 1071 1071 FT CARBOHYD 1072 1072 FT CARBOHYD 1073 1073 FT CARBOHYD 1074 1074 FT CARBOHYD 1075 1075 FT CARBOHYD 1076 1076 FT CARBOHYD 1077 1077 FT CARBOHYD 1078 1078 FT CARBOHYD 1079 1079 FT CARBOHYD 1080 1080 FT CARBOHYD 1081 1081 FT CARBOHYD 1082 1082 FT CARBOHYD 1083 1083 FT CARBOHYD 1084 1084 FT CARBOHYD 1085 1085 FT CARBOHYD 1086 1086 FT CARBOHYD 1087 1087 FT CARBOHYD 1088 1088 FT CARBOHYD 1089 1089 FT CARBOHYD 1090 1090 FT CARBOHYD 1091 1091 FT CARBOHYD 1092 1092 FT CARBOHYD 1093 1093 FT CARBOHYD 1094 1094 FT CARBOHYD 1095 1095 FT CARBOHYD 1096 1096 FT CARBOHYD 1097 1097 FT CARBOHYD 1098 1098 FT CARBOHYD 1099 1099 FT CARBOHYD 1100 1100 FT CARBOHYD 1101 1101 FT CARBOHYD 1102 1102 FT CARBOHYD 1103 1103 FT CARBOHYD 1104 1104 FT CARBOHYD 1105 1105 FT CARBOHYD 1106 1106 FT CARBOHYD 1107 1107 FT CARBOHYD 1108 1108 FT CARBOHYD 1109 1109 FT CARBOHYD 1110 1110 FT CARBOHYD 1111 1111 FT CARBOHYD 1112 1112 FT CARBOHYD 1113 1113 FT CARBOHYD 1114 1114 FT CARBOHYD 1115 1115 FT CARBOHYD 1116 1116 FT CARBOHYD 1117 1117 FT CARBOHYD 1118 1118 FT CARBOHYD 1119 1119 FT CARBOHYD 1120 1120 FT CARBOHYD 1121 1121 FT CARBOHYD 1122 1122 FT CARBOHYD 1123 1123 FT CARBOHYD 1124 1124 FT CARBOHYD 1125 1125 FT CARBOHYD 1126 1126 FT CARBOHYD 1127 1127 FT CARBOHYD 1128 1128 FT CARBOHYD 1129 1129 FT CARBOHYD 1130 1130 FT CARBOHYD 1131 1131 FT CARBOHYD 1132 1132 FT CARBOHYD 1133 1133 FT CARBOHYD 1134 1134 FT CARBOHYD 1135 1135 FT CARBOHYD 1136 1136 FT CARBOHYD 1137 1137 FT CARBOHYD 1138 1138 FT CARBOHYD 1139 1139 FT CARBOHYD 1140 1140 FT CARBOHYD 1141 1141 FT CARBOHYD 1142 1142 FT CARBOHYD 1143 1143 FT CARBOHYD 1144 1144 FT CARBOHYD 1145 1145 FT CARBOHYD 1146 1146 FT CARBOHYD 1147 1147 FT CARBOHYD 1148 1148 FT CARBOHYD 1149 1149 FT CARBOHYD 1150 1150 FT CARBOHYD 1151 1151 FT CARBOHYD 1152 1152 FT CARBOHYD 1153 1153 FT CARBOHYD 1154 1154 FT CARBOHYD 1155 1155 FT CARBOHYD 1156 1156 FT CARBOHYD 1157 1157 FT CARBOHYD 1158 1158 FT CARBOHYD 1159 1159 FT CARBOHYD 1160 1160 FT CARBOHYD 1161 1161 FT CARBOHYD 1162 1162 FT CARBOHYD 1163 1163 FT CARBOHYD 1164 1164 FT CARBOHYD 1165 1165 FT CARBOHYD 1166 1166 FT CARBOHYD 1167 1167 FT CARBOHYD 1168 1168 FT CARBOHYD 1169 1169 FT CARBOHYD 1170 1170 FT CARBOHYD 1171 1171 FT CARBOHYD 1172 1172 FT CARBOHYD 1173 1173 FT CARBOHYD 1174 1174 FT CARBOHYD 1175 1175 FT CARBOHYD 1176 1176 FT CARBOHYD 1177 1177 FT CARBOHYD 1178 1178 FT CARBOHYD 1179 1179 FT CARBOHYD 1180 1180 FT CARBOHYD 1181 1181 FT CARBOHYD 1182 1182 FT CARBOHYD 1183 1183 FT CARBOHYD 1184 1184 FT CARBOHYD 1185 1185 FT CARBOHYD 1186 1186 FT CARBOHYD 1187 1187 FT CARBOHYD 1188 1188 FT CARBOHYD 1189 1189 FT CARBOHYD 1190 1190 FT CARBOHYD 1191 1191 FT CARBOHYD 1192 1192 FT CARBOHYD 1193 1193 FT CARBOHYD 1194 1194 FT CARBOHYD 1195 1195 FT CARBOHYD 1196 1196 FT CARBOHYD 1197 1197 FT CARBOHYD 1198 1198 FT CARBOHYD 1199 1199 FT CARBOHYD 1200 1200 FT CARBOHYD 1201 1201 FT CARBOHYD 1202 1202 FT CARBOHYD 1203 1203 FT CARBOHYD 1204 1204 FT CARBOHYD 1205 1205 FT CARBOHYD 1206 1206 FT CARBOHYD 1207 1207 FT CARBOHYD 1208 1208 FT CARBOHYD 1209 1209 FT CARBOHYD 1210 1210 FT CARBOHYD 1211 1211 FT CARBOHYD 1212 1212 FT CARBOHYD 1213 1213 FT CARBOHYD 1214 1214 FT CARBOHYD 1215 1215 FT CARBOHYD 1216 1216 FT CARBOHYD 1217 1217 FT CARBOHYD 1218 1218 FT CARBOHYD 1219 1219 FT CARBOHYD 1220 1220 FT CARBOHYD 1221 1221 FT CARBOHYD 1222 1222 FT CARBOHYD 1223 1223 FT CARBOHYD 1224 1224 FT CARBOHYD 1225 1225 FT CARBOHYD 1226 1226 FT CARBOHYD 1227 1227 FT CARBOHYD 1228 1228 FT CARBOHYD 1229 1229 FT CARBOHYD 1230 1230 FT CARBOHYD 1231 1231 FT CARBOHYD 1232 1232 FT CARBOHYD 1233 1233 FT CARBOHYD 1234 1234 FT CARBOHYD 1235 1235 FT CARBOHYD 1236 1236 FT CARBOHYD 1237 1237 FT CARBOHYD 1238 1238 FT CARBOHYD 1239 1239 FT CARBOHYD 1240 1240 FT CARBOHYD 1241 1241 FT CARBOHYD 1242 1242 FT CARBOHYD 1243 1243 FT CARBOHYD 1244 1244 FT CARBOHYD 1245 1245 FT CARBOHYD 1246 1246 FT CARBOHYD 1247 1247 FT CARBOHYD 1248 1248 FT CARBOHYD 1249 1249 FT CARBOHYD 1250 1250 FT CARBOHYD 1251 1251 FT CARBOHYD 1252 1252 FT CARBOHYD 1253 1253 FT CARBOHYD 1254 1254 FT CARBOHYD 1255 1255 FT CARBOHYD 1256 1256 FT CARBOHYD 1257 1257 FT CARBOHYD 1258 1258 FT CARBOHYD 1259 1259 FT CARBOHYD 1260 1260 FT CARBOHYD 1261 1261 FT CARBOHYD 1262 1262 FT CARBOHYD 1263 1263 FT CARBOHYD 1264 1264 FT CARBOHYD 1265 1265 FT CARBOHYD 1266 1266 FT CARBOHYD 1267 1267 FT CARBOHYD 1268 1268 FT CARBOHYD 1269 1269 FT CARBOHYD 1270 1270 FT CARBOHYD 1271 1271 FT CARBOHYD 1272 1272 FT CARBOHYD 1273 1273 FT CARBOHYD 1274 1274 FT CARBOHYD 1275 1275 FT CARBOHYD 1276 1276 FT CARBOHYD 1277 1277 FT CARBOHYD 1278 1278 FT CARBOHYD 1279 1279 FT CARBOHYD 1280 1280 FT CARBOHYD 1281 1281 FT CARBOHYD 1282 1282 FT CARBOHYD 1283 1283 FT CARBOHYD 1284 1284 FT CARBOHYD 1285 1285 FT CARBOHYD 1286 1286 FT CARBOHYD 1287 1287 FT CARBOHYD 1288 1288 FT CARBOHYD 1289 1289 FT CARBOHYD 1290 1290 FT CARBOHYD 1291 1291 FT CARBOHYD 1292 1292 FT CARBOHYD 1293 1293 FT CARBOHYD 1294 1294 FT CARBOHYD 1295 1295 FT CARBOHYD 1296 1296 FT CARBOHYD 1297 1297 FT CARBOHYD 1298 1298 FT CARBOHYD 1299 1299 FT CARBOHYD 1300 1300 FT CARBOHYD 1301 1301 FT CARBOHYD 1302 1302 FT CARBOHYD 1303 1303 FT CARBOHYD 1304 1304 FT CARBOHYD 1305 1305 FT CARBOHYD 1306 1306 FT CARBOHYD 1307 1307 FT CARBOHYD 1308 1308 FT CARBOHYD 1309 1309 FT CARBOHYD 1310 1310 FT CARBOHYD 1311 1311 FT CARBOHYD 1312 1312 FT CARBOHYD 1313 1313 FT CARBOHYD 1314 1314 FT CARBOHYD 1315 1315 FT CARBOHYD 1316 1316 FT CARBOHYD 1317 1317 FT CARBOHYD 1318 1318 FT CARBOHYD 1319 1319 FT CARBOHYD 1320 1320 FT CARBOHYD 1321 1321 FT CARBOHYD 1322 1322 FT CARBOHYD 1323 1323 FT CARBOHYD 1324 1324 FT CARBOHYD 1325 1325 FT CARBOHYD 1326 1326 FT CARBOHYD 1327 1327 FT CARBOHYD 1328 1328 FT CARBOHYD 1329 1329 FT CARBOHYD 1330 1330 FT CARBOHYD 1331 1331 FT CARBOHYD 1332 1332 FT CARBOHYD 1333 1333 FT CARBOHYD 1334 1334 FT CARBOHYD 1335 1335 FT CARBOHYD 1336 1336 FT CARBOHYD 1337 1337 FT CARBOHYD 1338 1338 FT CARBOHYD 1339 1339 FT CARBOHYD 1340 1340 FT CARBOHYD 1341 1341 FT CARBOHYD 1342 1342 FT CARBOHYD 1343 1343 FT CARBOHYD 1344 1344 FT CARBOHYD 1345 1345 FT CARBOHYD 1346 1346 FT CARBOHYD 1347 1347 FT CARBOHYD 1348 1348 FT CARBOHYD 1349 1349 FT CARBOHYD 1350 1350 FT CARBOHYD 1351 1351 FT CARBOHYD 1352 1352 FT CARBOHYD 1353 1353 FT CARBOHYD 1354 1354 FT CARBOHYD 1355 1355 FT CARBOHYD 1356 1356 FT CARBOHYD 1357 1357 FT CARBOHYD 1358 1358 FT CARBOHYD 1359 1359 FT CARBOHYD 1360 1360 FT CARBOHYD 1361 1361 FT CARBOHYD 1362 1362 FT CARBOHYD 1363 1363 FT CARBOHYD 1364 1364 FT CARBOHYD 1365 1365 FT CARBOHYD 1366 1366 FT CARBOHYD 1367 1367 FT CARBOHYD 1368 1368 FT CARBOHYD 1369 1369 FT CARBOHYD 1370 1370 FT CARBOHYD 1371 1371 FT CARBOHYD 1372 1372 FT CARBOHYD 1373 1373 FT CARBOHYD 1374 1374 FT CARBOHYD 1375 1375 FT CARBOHYD 1376 1376 FT CARBOHYD 1377 1377 FT CARBOHYD 1378 1378 FT CARBOHYD 1379 1379 FT CARBOHYD 1380 1380 FT CARBOHYD 1381 1381 FT CARBOHYD 1382 1382 FT CARBOHYD 1383 1383 FT CARBOHYD 1384 1384 FT CARBOHYD 1385 1385 FT CARBOHYD 1386 1386 FT CARBOHYD 1387 1387 FT CARBOHYD 1388 1388 FT CARBOHYD 1389 1389 FT CARBOHYD 1390 1390 FT CARBOHYD 1391 1391 FT CARBOHYD 1392 1392 FT CARBOHYD 1393 1393 FT CARBOHYD 1394 1394 FT CARBOHYD 1395 1395 FT CARBOHYD 1396 1396 FT CARBOHYD 1397 1397 FT CARBOHYD 1398 1398 FT CARBOHYD 1399 1399 FT CARBOHYD 1400 1400 FT CARBOHYD 1401 1401 FT CARBOHYD 1402 1402 FT CARBOHYD 1403 1403 FT CARBOHYD 1404 1404 FT CARBOHYD 1405 1405 FT CARBOHYD 1406 1406 FT CARBOHYD 1407 1407 FT CARBOHYD 1408 1408 FT CARBOHYD 1409 1409 FT CARBOHYD 1410 1410 FT CARBOHYD 1411 1411 FT CARBOHYD 1412 1412 FT CARBOHYD 1413 1413 FT CARBOHYD 1414 1414 FT CARBOHYD 1415 1415 FT CARBOHYD 1416 1416 FT CARBOHYD 1417 1417 FT CARBOHYD 1418 1418 FT CARBOHYD 1419 1419 FT CARBOHYD 1420 1420 FT CARBOHYD 1421 1421 FT CARBOHYD 1422 1422 FT CARBOHYD 1423 1423 FT CARBOHYD 1424 1424 FT CARBOHYD 1425 1425 FT CARBOHYD 1426 1426 FT CARBOHYD 1427

```

RESULT /  

ITB5_PAPCY  

ID ITB5_PAPCY  

AC Q01411;  

STANDARD;  

PRT; 655 AA.

```

SEQUENCE G-F-P-G-E-R IN COLLAGEN. INTEGRINS ALPHA-2/BETA-1, ALPHA-3/BETA-1, ALPHA-4/BETA-1, ALPHA-5/BETA-1, ALPHA-6/BETA-1, ALPHA-8/BETA-1, ALPHA-10/BETA-1, ALPHA-11/BETA-1 AND ALPHA-V/BETA-1 ARE RECEPTORS FOR FIBRONECTIN. ALPHA-4/BETA-1 RECOGNIZES ONE OR MORE DOMAINS WITHIN THE ALTERNATIVELY SPliced CS-1 AND CS-5 REGIONS OF FIBRONECTIN. INTEGRIN ALPHA-5/BETA-1 IS A RECEPTOR FOR FIBRINOGEN. INTEGRIN ALPHA-1/BETA-1, ALPHA-2/BETA-1, ALPHA-6/BETA-1 AND ALPHA-9/BETA-1 ARE RECEPTORS FOR VCAM1 AND RECOGNIZES THE SEQUENCE Q-T-D-S IN VCAM1. INTEGRIN ALPHA-9/BETA-1 IS A RECEPTOR FOR VCAM1, CYTOKIN AND OSTEOPONTIN. IT RECOGNIZES THE SEQUENCE A-E-I-D-G-I-E-L IN CYTOKIN. INTEGRIN ALPHA-3/BETA-1 IS A RECEPTOR FOR EPILIGRIN AND THROMbospondin. INTEGRIN ALPHA-V/BETA-1 IS A RECEPTOR FOR VITROKIN. BETA-1 INTEGRINS RECOGNIZE THE SEQUENCE R-G-D IN A WIDE ARRAY OF LIGANDS.

- SUBUNIT: HETEROdIMER OF AN ALPHA AND A BETA SUBUNIT. BETA-1 ASSOCIATES WITH EITHER ALPHA-1, ALPHA-2, ALPHA-3, ALPHA-4, ALPHA-5, ALPHA-6, ALPHA-7, ALPHA-8, ALPHA-9, ALPHA-10, ALPHA-11 OR ALPHA-V.

- SUBCELLULAR LOCATION: TYPE I membrane protein.

- SIMILARITY: BELONGS TO THE INTEGRIN BETA CHAIN FAMILY.

- SIMILARITY: CONTAINS 1 VWFa-LIKE DOMAIN.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation, the European Bioinformatics Institute. There are no restrictions on use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commerce entities requires a license agreement (See <http://www.isb-sib.ch/announcements.html>).

CC EMBL: M14049; AAA48926.1; -.

CC PIR: A23947; IUC3.

CC DR InterPro; IPR000561; EGF-like.

CC DR InterPro; IPR02369; Integrin_B.

CC DR InterPro; IPR001169; Integrin_beta_C.

CC DR InterPro; IPR003659; PSI.

CC DR InterPro; IPR02035; VWFa.

CC DR Pfam; PF00362; Integrin_B; 1.

CC DR PRINTS; PRO0186; INTEGRIN_B.

CC DR PRODOM; PRO0811; Integrin_B; 1.

CC DR SMART; SM00187; INB; 1.

CC DR SMART; SM00423; PSI; 1.

CC DR PROSITE; PS00243; INTEGRIN_BETA_3.

CC DR PROSITE; PS00242; EGF_1; UNKNOWN.

CC KW Integrin; Cell adhesion; Receptor; Transmembrane; Glycoprotein; Repeat; Phosphorylation; Signal.

CC KW SIGNAL 1 24 BY SIMILARITY.

CC FT CHAIN 25 803 INTEGRIN BETA-1.

CC FT DOMAIN 25 733 EXTRACELLULAR (POTENTIAL).

CC FT TRANSMEM 734 756 POTENTIAL.

CC FT DOMAIN 144 803 Cytosolic (POTENTIAL).

CC FT DOMAIN 471 520 VWFa-LIKE.

CC FT REPEAT 471 520 CYSTEINE-RICH TANDEM REPEATS.

CC FT REPEAT 521 564 I.

CC FT REPEAT 521 603 II.

CC FT REPEAT 565 603 III.

CC FT REPEAT 604 640 IV.

CC FT DISULFID 31 469 BY SIMILARITY.

CC FT DISULFID 39 449 BY SIMILARITY.

CC FT DISULFID 42 79 BY SIMILARITY.

CC FT DISULFID 52 68 BY SIMILARITY.

CC FT DISULFID 211 217 BY SIMILARITY.

CC FT DISULFID 265 305 BY SIMILARITY.

CC FT DISULFID 405 419 BY SIMILARITY.

CC FT DISULFID 439 696 BY SIMILARITY.

CC FT DISULFID 467 471 BY SIMILARITY.

CC FT DISULFID 482 494 BY SIMILARITY.

CC FT DISULFID 491 530 BY SIMILARITY.

CC FT DISULFID 496 505 BY SIMILARITY.

CC FT DISULFID 507 521 BY SIMILARITY.

CC FT DISULFID 536 541 BY SIMILARITY.

RESULT 8

ITBL-CHICK STANDARD: PRT: 803 AA.

ID ITBL-CHICK

AC P07250; (Rel. 07, Created)

DT 01-APR-1988 (Rel. 07, Last sequence update)

DT 01-MAR-2002 (Rel. 41, Last annotation update)

DE Integrin beta-1 precursor (CSAT antigen) (JG22 antigen) (RGD-receptor).

DE ITGB1

OS Gallus gallus (Chicken).

OC Bivalvia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauia; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.

OC NCBI_Taxid:9031;

[1] SEQUENCE FROM N.A.

RP TISSUE=Embryonic fibroblast;

RC MEDLINE=86245073; PubMed=348786;

RA Tamkin J.W., Desimone D.W., Fonda D., Patel R.S., Buck C., Horwitz A.F., Hynes R.O.

RT "Structure of integrin, a glycoprotein involved in the transmembrane linkage between fibronectin and actin.";

RT Cell 46:271-282 (1986).

RT -!- FUNCTION: INTEGRINS ALPHA-1/BETA-1, ALPHA-2/BETA-1, ALPHA-10/BETA-1 AND ALPHA-11/BETA-1 ARE RECEPTORS FOR COLLAGEN. INTEGRINS ALPHA-1/BETA-1 AND ALPHA-2/BETA-2 RECOGNIZE THE PROLINE-HYDROXYLATED

PT DISURFD 538 BY SIMILARITY.
 PT DISURFD 543 BY SIMILARITY.
 PT DISURFD 560 BY SIMILARITY.
 PT DISURFD 579 BY SIMILARITY.
 PT DISURFD 581 BY SIMILARITY.
 PT DISURFD 586 BY SIMILARITY.
 PT DISURFD 597 BY SIMILARITY.
 PT DISURFD 618 BY SIMILARITY.
 PT DISURFD 620 BY SIMILARITY.
 PT DISURFD 625 BY SIMILARITY.
 PT DISURFD 635 BY SIMILARITY.
 PT DISURFD 645 BY SIMILARITY.
 PT DISURFD 651 BY SIMILARITY.
 PT DISURFD 670 BY SIMILARITY.
 MOD_RES 25 BLOCKED.
 MOD_RES 788 PHOSPHORYLATION (BY TYR-KINASES)
 (POTENTIAL).
 FT CARBOHYD 216 (POTENTIAL).
 FT CARBOHYD 273 (POTENTIAL).
 FT CARBOHYD 367 (POTENTIAL).
 FT CARBOHYD 410 (POTENTIAL).
 FT CARBOHYD 421 (POTENTIAL).
 FT CARBOHYD 433 (POTENTIAL).
 FT CARBOHYD 445 (POTENTIAL).
 FT CARBOHYD 486 (POTENTIAL).
 FT CARBOHYD 525 (POTENTIAL).
 FT CARBOHYD 589 (POTENTIAL).
 FT CARBOHYD 624 (POTENTIAL).
 FT CARBOHYD 674 (POTENTIAL).
 SEQUENCE 803 AA: 88553 MN: 2F6FERCDF2C80457 CRC64;

Query Match 44.5%; Score 1846.5; DB 1; Length 803;
 Best Local Similarity 45.1%; Pred: No 1.1e-17;
 Matches 353; Conservative 136; Mismatches 266; Indels 27; Gaps 15;

QY 1 GPNICITTRGVSSQCLAVSPMWC-SDEALPLGSP--RDLKENLKDQKAPESIEF 56
 | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
 Db 27 GGSPECIKANAKSCKCCTGAGPNCWCKTDPDQEGEPISARCDLAAKSKCGPEQDIE 86
 | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
 Qy 57 PVSARVELDRPSPDKGSDSSO---VTOVSPOINTALRPLPDDSKNSFQVQVEDYP 111
 | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
 Db 87 PROSKYLEDREYETNRKIGAAEKLKPEATQDOPQKYLQRLRQPEOTFSKFERADYP 146
 | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
 Qy 112 VDTIVYLMOLSYSKDDMSIONGKTLQTMQRMLTSNLRIGRAFDKPVSPWYISPE 171
 | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
 Db 147 IDLYYLMOLSYSKDDLENVKSLGITALMREMKITSDFRIGGSFVERTVMPYIISTPK 206
 | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
 Qy 172 ALNPICYDMKTCILPMEGKYKHLTLDVTRNEEVKQSYISRNRAPEGEDAIMQATV 231
 | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
 Db 207 -LNPCTG-DONOTSPSYKNVLSLTSGNKFNELVKGQHISGNDNHSASTMDYPS 291
 | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
 Qy 232 CDPKIGRNNDASHLVLFTDATHTIALDGLLAGIWPNDQGCHVGWSNDHISASTMDYPS 322
 | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
 Db CGQIGKWN-VTLLVSTDAFGHAGDKGKLGIVAPDGKHL-ENNYMYSHYDPS 322
 | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
 292 LGMTKEKISOKXINLIFAVTENVNLQYNSSELLPGTVGULSMDSNVQQLIVDAYGKI 351
 | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
 Db 323 IAHWVKOLESENNTIOTIATVTEEEQAYKELKMLIPSAVGLLSSNSNVIOLIADYNSL 382
 | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
 Qy 352 RSVVELEVDRDLPEELSLSFNATCLN--NEVTPGKSCGMGLTGTWFSSTEAKVKGCPQE 409
 | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
 Db 383 SSEVILENSKLFKMEVTSYKCKNGVNDTQEDGRKCSNSISIGDEVRFELINTNECPKK 442
 | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
 Qy 410 -KKSFTKPKVERFKDSLIVQWVDFDCACQACQAEPRSHRCNGTFFECGVCRCPGWL 468
 | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
 Db 443 GQNETIKKPKLGFTEVEIHIQFCIDCICQCBEGPNSPACIDNGTFFECGACRCHBRIG 502
 | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
 Db 469 SOCECSEEDYRQSQDE-CSPREQQPVCQSRGECLCQGQCVHSDEGK-ITGKCBCCD 525
 | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
 Db 503 RLCECSTDEVNSEDMDAYCRENSTEICCSNNGCCTGQCVCKKRENTNEVYSGKCECDN 562
 | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
 Qy 526 FSCVRYKEMCSGGHGOQSCGCGCLCSDDWGTYCNCITRTDQMMSSNGLCGRGKCECGS 585

QY 642 VKEL-KDTGKDAV-NCTYKKNEDDCVVRQYEDSSKSILVVEERECPPKGPDILVLLS 699
 | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
 Db 681 RKLQPQVHPDPLSHCKEKGDCWVFTSYNSNGERSAHSVVEPECPSPGDPDIPIVAG 740
 | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
 Qy 700 VMAAILLIGLAALITWLLITHDREFAKEEERAKRDTANRPLYKATSTENITV 759
 | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
 Db 741 VVAGIVLIGLALLIWLMIHDRREFAKEREMNAKWDJGENPIYKAVTTVNPKY 800
 | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
 Qy 760 RG 761
 Db 801 EG 802
 RA
 RQ
 RN
 RP
 RA
 RQ
 RN
 RP
 SEQUENCE FROM N. A.
 Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.
 -1- FUNCTION: INTEGRINS ALPHA-1/BETA-1, ALPHA-2/BETA-1, ALPHA-10/BETA-1 AND ALPHA-11/BETA-1 ARE RECEPTORS FOR COLLAGEN. INTEGRINS ALPHA-1/BETA-1 AND ALPHA-2/BETA-2 RECOGNIZE THE PROLINE-HYDROXYLATED SEQUENCE G-F-P-G-E-R IN COLLAGEN. INTEGRINS ALPHA-2/BETA-1, ALPHA-3/BETA-1, ALPHA-4/BETA-1, ALPHA-5/BETA-1, ALPHA-8/BETA-1, ALPHA-10/BETA-1, ALPHA-11/BETA-1 AND ALPHA-V/BETA-1 ARE RECEPTORS FOR FIBRONECTIN. INTEGRIN ALPHA-4/BETA-1 RECOGNIZES ONE OR MORE DOMAINS WITHIN THE ALTERNATIVELY SPLICED CS-1 AND CS-5 REGIONS OF FIBRONECTIN. INTEGRIN ALPHA-5/BETA-1 IS A RECEPTOR FOR FIBRONECTIN. INTEGRIN ALPHA-1/BETA-1, ALPHA-2/BETA-1, ALPHA-4/BETA-1, ALPHA-5/BETA-1 AND ALPHA-7/BETA-1 ARE RECEPTORS FOR LAMININ. INTEGRIN ALPHA-4/BETA-1 IS A RECEPTOR FOR VCAN1 AND RECOGNIZES THE SEQUENCE Q-T-D-S IN VCAN1. INTEGRIN ALPHA-9/BETA-1 IS A RECEPTOR FOR VCAN1, CYTOKININ AND OSTEOPONTIN. IT RECOGNIZES THE SEQUENCE A-E-T-D-G-I-E-L IN CYTOKININ. INTEGRIN ALPHA-3/BETA-1 IS A RECEPTOR FOR EPILIGRIN AND THROMBOSPONDIN. INTEGRIN ALPHA-V/BETA-1 IS A RECEPTOR FOR VITRONECTIN. BETA-1 INTEGRINS RECOGNIZE THE SEQUENCE R-G-D IN A WIDE ARRAY OF LIGANDS.
 -1- SUBUNIT: HETEROODIMER OF AN ALPHA AND A BETA SUBUNIT. BETA-1 ASSOCIATES WITH EITHER ALPHA-1, ALPHA-2, ALPHA-3, ALPHA-4, ALPHA-5, ALPHA-6, ALPHA-7, ALPHA-8, ALPHA-9, ALPHA-10, ALPHA-11 OR ALPHA-V.
 -1- SUBCELLULAR LOCATION: TYPE I membrane protein.
 -1- SIMILARITY: BELONGS TO THE INTEGRIN BETA CHAIN FAMILY.
 -1- SIMILARITY: CONTAINS 1 VWA-LIKE DOMAIN.
 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation in the European Bioinformatics Institute. There are no restrictions on its use by non profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial use requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 DR EMBL_U27351; AAC19407.1; -

Db 563 FNCDRSNLICGGNGLICKRCVCECFPNTGSAACDCSILDTPTCAGNGQICNGRGTCECGT 622
 Qy 586 CYCIOQPSYGDCEKCTCPACTFKKECKFDRGALHDENTCNRIC---RDEIES 641
 | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
 Db 623 CINCQDPKQGPQCEMCQTCPLGQVCAEIKDCVOCRAEKG-EKKEVCSQECMENTRVS 680
 Qy 642 VKEL-KDTGKDAV-NCTYKKNEDDCVVRQYEDSSKSILVVEERECPPKGPDILVLLS 699
 | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
 Db 681 RKLQPQVHPDPLSHCKEKGDCWVFTSYNSNGERSAHSVVEPECPSPGDPDIPIVAG 740
 | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
 Qy 700 VMAAILLIGLAALITWLLITHDREFAKEEERAKRDTANRPLYKATSTENITV 759
 | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
 Db 741 VVAGIVLIGLALLIWLMIHDRREFAKEREMNAKWDJGENPIYKAVTTVNPKY 800
 | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
 Qy 760 RG 761
 Db 801 EG 802
 RA
 RQ
 RN
 RP
 SEQUENCE FROM N. A.
 Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.
 -1- FUNCTION: INTEGRINS ALPHA-1/BETA-1, ALPHA-2/BETA-1, ALPHA-10/BETA-1 AND ALPHA-11/BETA-1 ARE RECEPTORS FOR COLLAGEN. INTEGRINS ALPHA-1/BETA-1 AND ALPHA-2/BETA-2 RECOGNIZE THE PROLINE-HYDROXYLATED SEQUENCE G-F-P-G-E-R IN COLLAGEN. INTEGRINS ALPHA-2/BETA-1, ALPHA-3/BETA-1, ALPHA-4/BETA-1, ALPHA-5/BETA-1, ALPHA-8/BETA-1, ALPHA-10/BETA-1, ALPHA-11/BETA-1 AND ALPHA-V/BETA-1 ARE RECEPTORS FOR FIBRONECTIN. INTEGRIN ALPHA-4/BETA-1 RECOGNIZES ONE OR MORE DOMAINS WITHIN THE ALTERNATIVELY SPLICED CS-1 AND CS-5 REGIONS OF FIBRONECTIN. INTEGRIN ALPHA-5/BETA-1 IS A RECEPTOR FOR FIBRONECTIN. INTEGRIN ALPHA-1/BETA-1, ALPHA-2/BETA-1, ALPHA-4/BETA-1, ALPHA-5/BETA-1 AND ALPHA-7/BETA-1 ARE RECEPTORS FOR LAMININ. INTEGRIN ALPHA-4/BETA-1 IS A RECEPTOR FOR VCAN1 AND RECOGNIZES THE SEQUENCE Q-T-D-S IN VCAN1. INTEGRIN ALPHA-9/BETA-1 IS A RECEPTOR FOR VCAN1, CYTOKININ AND OSTEOPONTIN. IT RECOGNIZES THE SEQUENCE A-E-T-D-G-I-E-L IN CYTOKININ. INTEGRIN ALPHA-3/BETA-1 IS A RECEPTOR FOR EPILIGRIN AND THROMBOSPONDIN. INTEGRIN ALPHA-V/BETA-1 IS A RECEPTOR FOR VITRONECTIN. BETA-1 INTEGRINS RECOGNIZE THE SEQUENCE R-G-D IN A WIDE ARRAY OF LIGANDS.
 -1- SUBUNIT: HETEROODIMER OF AN ALPHA AND A BETA SUBUNIT. BETA-1 ASSOCIATES WITH EITHER ALPHA-1, ALPHA-2, ALPHA-3, ALPHA-4, ALPHA-5, ALPHA-6, ALPHA-7, ALPHA-8, ALPHA-9, ALPHA-10, ALPHA-11 OR ALPHA-V.
 -1- SUBCELLULAR LOCATION: TYPE I membrane protein.
 -1- SIMILARITY: BELONGS TO THE INTEGRIN BETA CHAIN FAMILY.
 -1- SIMILARITY: CONTAINS 1 VWA-LIKE DOMAIN.
 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation in the European Bioinformatics Institute. There are no restrictions on its use by non profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial use requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 DR EMBL_U27351; AAC19407.1; -

Query	Match	Score	Length	DB	Matches	Conservative	Mismatches	Indels	Gaps	13;	
Best Local Similarity		44.1%	798	DB1	44.6%	Score 1833;	DB 1;	Length 798;	pred. No. 8.7e-17;		
PROSTATE; PS000243;	PROSTATE; PS000243;	44.1%	798	Db	348	3	3	28	13;		
DR	DR	44.1%	798	QY	3	NICITRGSSCQCLAVSPMCACSD---EALPLGSPRDIKENLKLKDCAPESIEFP	57	268			
InterPro; IPR002369; Integrin_B.	InterPro; IPR002369; Integrin_B.	44.1%	798	Db	25	-	-	-	-	57	
DR	DR	44.1%	798	QY	-	-	-	-	-	57	
Interrr; IPR001169; Integrin_beta_C.	Interrr; IPR001169; Integrin_beta_C.	44.1%	798	Db	25	NRCUAKANAKSCGCIOAQPNCGWCVNSTFLOEGMPT-SARCDDEALKKGKCHPPDDEINP	83	268			
DR	DR	44.1%	798	QY	-	-	-	-	-	83	
InterPro; IPR002035; vWFA.	InterPro; IPR002035; vWFA.	44.1%	798	Db	58	VSEARVLDRPLSDKGSDSSQ---WTQVSPORALRIRPDDSKNF1QVRQVEDY	PV	112			
PRam; PF00362; Integrin_B; 1.	PRam; PF00362; Integrin_B; 1.	44.1%	798	QY	-	-	-	-	-	112	
DR	DR	44.1%	798	Db	84	RGSKDVKKNKNVNRSKGTAEKOQEDITQOPOOLVQLQRLSGEPOTFLFKRRADEY	P	143			
prod0; PRO001811; Integrin_B; 1.	prod0; PRO001811; Integrin_B; 1.	44.1%	798	QY	-	-	-	-	-	143	
DR	DR	44.1%	798	QY	-	-	-	-	-	143	
SMART; SM00001; EGF_like; 1.	SMART; SM00001; EGF_like; 1.	44.1%	798	Db	SMART	SM000423; PSI; 1.					
DR	DR	44.1%	798	QY	-	-	-	-	-		
SMART; SM00327; vWA; 1.	SMART; SM00327; vWA; 1.	44.1%	798	Db	SMART	SM00423; PSI; 1.					
DR	DR	44.1%	798	QY	-	-	-	-	-		
PROSTATE; PS000243;	PROSTATE; PS000243;	44.1%	798	Db	SMART	SM00423; PSI; 1.					
DR	DR	44.1%	798	QY	-	-	-	-	-		
Integrin; Cell adhesion; Receptor; Transmembrane; Glycoprotein; Repeat; Signal; Phosphorylation.	Integrin; Cell adhesion; Receptor; Transmembrane; Glycoprotein; Repeat; Signal; Phosphorylation.	44.1%	798	Db	203	LRNPFC-TSEQNCTSPFSKVNVLTDKEVFNELVKGKRSIGNLSDPEGGFDAMQAVC	261	203			
FT	FT	44.1%	798	QY	-	-	-	-	-	261	
CHAIN	CHAIN	44.1%	798	Db	233	DEKIGWQRDASHLVLFTDAKTHIALQGRLAGIVQWPNQGCVGNSHNHYSTMDYPSL	292	233			
FT	FT	44.1%	798	QY	-	-	-	-	-	292	
DOMAIN	DOMAIN	44.1%	798	Db	262	GSLIGWGN-VRLLIVFSTDAGHFAGDKRGGIQLVPMQCHLEND-VYTMHYDYSI	319	262			
FT	FT	44.1%	798	QY	-	-	-	-	-	319	
DOMAIN	DOMAIN	44.1%	798	Db	380	SEVILENSKLPKEGVITSYKSYKNGVNGTGENGRCGCNSISIGDEVQFELSTANKPNK	439	380			
FT	FT	44.1%	798	QY	-	-	-	-	-	439	
REPEAT	REPEAT	44.1%	798	Db	440	SEITKIKPLGFTEVEILQFICECECONEGIPSSPKCHEGNSFECGACRNEGRVGRH	499	440			
FT	FT	44.1%	798	QY	-	-	-	-	-	499	
DISULFID	DISULFID	44.1%	798	Db	471	CEC5EEDRPSQDDE-CPPREGGPPVCSORGECGECIICGCVCHSSDFGK--ITGKCECDDFS	527	471			
FT	FT	44.1%	798	QY	-	-	-	-	-	527	
DISULFID	DISULFID	44.1%	798	Db	500	CECSTDDEVNSEDMDAYCKRKENNSEBICSRNGECVGQCVORKRQNTNLISKGKFCEDFN	559	500			
FT	FT	44.1%	798	QY	-	-	-	-	-	559	
DISULFID	DISULFID	44.1%	798	Db	528	CYRKGEMSGHGCSCDCLCSDWGYYCNCNTTDTGMSNGLGSGRKCEGSCV	587	528			
FT	FT	44.1%	798	QY	-	-	-	-	-	587	
DISULFID	DISULFID	44.1%	798	Db	560	CDRSGNLIGGGNGVCKRYCVCNECPYNTOSACDCSLSDTSCMATINGQICNGRICEGACK	619	560			
FT	FT	44.1%	798	QY	-	-	-	-	-	619	
DISULFID	DISULFID	44.1%	798	Db	588	CIQPSIGKGTCKECPKACTKRECKKDKRHDIDENICRGR--DELESVK	643	588			
FT	FT	44.1%	798	QY	-	-	-	-	-	643	
DISULFID	DISULFID	44.1%	798	Db	620	CTDPKFQGPICMCMQTCGCVBKEVQCAFNCNPKGERD-TCAQCSHFNITKVNRD	677	620			
FT	FT	44.1%	798	QY	-	-	-	-	-	677	
DISULFID	DISULFID	44.1%	798	Db	644	ELKDQGK--DAVNCYKINEDDCVVRFQYEDSSGKSTLWVERPECPKGPDLVLVLSVM	701	644			
FT	FT	44.1%	798	QY	-	-	-	-	-	701	
DISULFID	DISULFID	44.1%	798	Db	678	KLPQFGQVPLSKRKEKWDCCMRYFTTSVNGNEAIWVWPECPGDPITVAGWV	737	678			
FT	FT	44.1%	798	QY	-	-	-	-	-	737	
DISULFID	DISULFID	44.1%	798	Db	702	GAIIILIGLALLTILKLLTIDRERAFEEERAKRDNTANPPLYKATSTFTNTYRG	761	702			
FT	FT	44.1%	798	QY	-	-	-	-	-	761	
DISULFID	DISULFID	44.1%	798	Db	738	AGIVLIGLALLIWLKLLMIHDREFAKERKMNKWDIGENPIYKAVTVNPKEG	797	738			
FT	FT	44.1%	798	QY	-	-	-	-	-	797	
DISULFID	DISULFID	44.1%	798	Db	798	AA					
FT	FT	44.1%	798	QY	-	-	-	-	-		
MOD_RES	MOD_RES	44.1%	798	Db	RESULT	10					
CARBOHYD	CARBOHYD	50	50	ITBL_HUMAN	ITBL_HUMAN	STANDARD	PRT	798	AA		
CARBOHYD	CARBOHYD	94	94	ITBL_HUMAN	ITBL_HUMAN	STANDARD	PRT	798	AA		
CARBOHYD	CARBOHYD	97	97	AC	P05556; P78466; P78467; Q13089; Q14647; Q13090; Q13212; Q13091;						
CARBOHYD	CARBOHYD	212	212	AC	Q14622;						
CARBOHYD	CARBOHYD	269	269	DT	01-NOV-1998 (Rel. 09, Created)						
CARBOHYD	CARBOHYD	363	363	DT	01-NOV-1998 (Rel. 09, Last sequence update)						
CARBOHYD	CARBOHYD	406	406	DT	01-MAR-2002 (Rel. 41, Last annotation update)						
CARBOHYD	CARBOHYD	417	417	DE	Integrin beta-1 precursor (Fibronectin receptor beta subunit)						
CARBOHYD	CARBOHYD	481	481	DE	(CD29) (Integrin VLA-4 beta subunit)						
CARBOHYD	CARBOHYD	520	520	GN	ITGB1 OR FNBR.						
CARBOHYD	CARBOHYD	584	584	OS	Homo sapiens (Human).						
CARBOHYD	CARBOHYD	669	669	OC	Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;						
SEQUENCE	SEQUENCE	798	AA	OC	Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.						
FT	FT	798	AA	NCBI_TaxID=9606;							
FT	FT	798	AA	NCBI_TaxID=9606;							
RC	RC	798	AA	NCBI_TaxID=9606;							
Query	Match	Score	Length	DB	SEQUENCE FROM N.A. (ISOFORM BETA-1A).						
Best Local Similarity		44.1%	798	RC	TISSUE-Placenta;						

RX MEDLINE=88007843; PubMed=298481; CC

RX Argraves W.S., Suzuki S., Arai H., Thompson K., Pierschbacher M.D., CC

RA Ruoslahti E., "Amino acid sequence of the human fibronectin receptor.," CC

RT [2] Cell Biol. 105:1183-1190(1987). CC

RN RX SEQUENCE OF 717-757 FROM N.A., AND CHARACTERIZATION OF BETA-1B. CC

RX MEDLINE=3320984; PubMed=7631433; CC

RA Balzac F., Belkin A.M., Koteliansky V.E., Balabanov Y.V., Altruda F., CC

RA Silengo L., Tarone G.; "Expression and functional analysis of a cytoplasmic domain variant of CC

RT the beta 1 integrin subunit.," J. Cell Biol. 121:171-178(1993). CC

RN [3] RX SEQUENCE OF 717-757 FROM N.A., AND FUNCTION. CC

RX MEDLINE=95014744; PubMed=7533423; CC

RA Balzac F., Reitta S.F., Albini A., Melchiorri A., Koteliansky V.E., CC

RA Geuna M., Silengo L., Tarone G.; "Expression of beta 1B integrin isoform in CHO cells results in a CC

RT dominant negative effect on cell adhesion and motility.," J. Cell Biol. 127:557-565(1994). CC

RX PARTIAL SEQUENCE FROM N.A. (ISOFORMS BETA-1A; BETA-1B; BETA-1C AND CC

RP BETA-1D). CC

RC TISSUE=Skeletal muscle; MEDLINE=95398646; PubMed=7745396; CC

RX "Novel isoform of beta 1 integrin expressed in skeletal and cardiac CC

RT muscle.," Biochem. Biophys. Res. Commun. 214:279-285(1995). CC

RN [5] RX PARTIAL SEQUENCE FROM N.A. (ISOFORM BETA-1B). CC

RC TISSUE=Skeletal muscle; MEDLINE=9105539; PubMed=2249781; CC

RA Altruda F., Cervella P., Botta C., Botta G., Stefanuto G., CC

RA Silengo L.; "A human integrin beta 1 subunit with a unique cytoplasmic domain CC

RT generated by alternative mRNA processing.," Gene 95:261-266(1990). CC

RN [6] RX PARTIAL SEQUENCE FROM N.A. (ISOFORM BETA-1D). CC

RC MEDLINE=95377431; PubMed=7544298; CC

RA van der Flier A., Kuiken T., Baudoin C., van der Neut R., CC

RA Sonnenberg A.; "A novel beta 1 integrin isoform produced by alternative splicing: CC

RT unique expression in cardiac and skeletal muscle.," FEBS Lett. 369:340-344(1995). CC

RN [7] RX PARTIAL SEQUENCE FROM N.A. (ISOFORMS BETA-1A AND BETA-1C). CC

RC TISSUE=Cervical carcinoma; MEDLINE=92202279; PubMed=1551917; CC

RA Languino L.R., Ruoslahti E.; "An alternative form of the integrin beta 1 subunit with a variant CC

RT cytoplasmic domain.," J. Biol. Chem. 267:7116-7120(1992). CC

RN [8] RX PARTIAL SEQUENCE FROM N.A. (ISOFORM BETA-1C-2). CC

RC MEDLINE=98161805; PubMed=494044; CC

RA Svineng G., Faesler R., Johansson S.; "Identification of beta1C-2, a novel variant of the integrin beta 1 subunit generated by utilization of an alternative splice acceptor site in exon C.," Biochem. J. 330:1255-1263(1998). CC

RT - FUNCTION: INTEGRINS ALPHA-1/BETA-1, ALPHA-2/BETA-1, ALPHA-10/BETA-1, AND ALPHA-11/BETA-1 ARE RECEPTORS FOR COLLAGEN. INTEGRINS ALPHA-1/BETA-1 AND ALPHA-2/BETA-2 RECOGNIZE THE PROLINE-HYDROXYLATED SEQUENCE G-F-P-G-E-R IN COLLAGEN. INTEGRINS ALPHA-2/BETA-1, ALPHA-3/BETA-1, ALPHA-4/BETA-1, ALPHA-5/BETA-1, ALPHA-8/BETA-1, ALPHA-10/BETA-1, AND ALPHA-11/BETA-1 ARE RECEPTORS FOR FIBRONECTIN. INTEGRINS ALPHA-4/BETA-1 RECOGNIZES ONE OR MORE DOMAINS WITHIN THE ALTERNATIVELY SPliced CS-1 AND CS-5 REGIONS OF FIBRONECTIN. INTEGRIN ALPHA-5/BETA-1 IS A RECEPTOR FOR FIBRONECTIN. INTEGRIN ALPHA-1/BETA-1, ALPHA-2/BETA-1, ALPHA-6/BETA-1 AND ALPHA-7/BETA-1. SMART; SM00187; INB; 1. SMART; SM0023; PSI; 1.

DR	SMART: SM00327; VWA: 1.	Db	144 LYVMDLSFSKMDLLENVYKSLGALMTEKIKTSDFRIGFSIWEKTVNPYISTTPAKL 203
DR	PROSITE: PS00243; INTEGRIN_BETA: 3.	Qy	174 ENPCYDMKTCIPLPMGKYKIVLUTDQVTRNEENVKQSVRSRNDAPECGFDAMQATVCD 233
KW	Integrin; Cell adhesion; Receptor; Transmembrane; Glycoprotein;	Db	204 -NPCIN-DONCTSFSYKAVNLTKDGKLFNDLVKGKQOISGNLSDPECGFDAMQAVCG 261
Repeat; Signal; Phosphorylation.	FT SIGNAL 1 21	Qy	234 EKIGWRNDASHLHVFTTDKTHFALDGRGAGIVQPNDCQHVGSDNHSASITMDYSLG 293
FT	CHAIN 22 798	Db	262 EQIGRN-VTRLYFVSTDAGFHAGDGKIGKIVLNDGRCHL-HGNMVTMHSYDPSIA 319
FT	DOMAIN 728 751	Qy	294 LMTEKLSONNLUFAWENNVNVQNYSELINGCTVGLSMUSSNVQIQLTDAYGKRS 353
FT	DOMAIN 752 798	Db	320 HLWQLSENNQIQTFAVTFDFOPVQELKRNLPKSAVETLSSNSNVYQIOLIIDSYNSLSS 379
FT	DOMAIN 139 377	Qy	354 KVELEVRDIEFEELSLISFNATCLNNNEVIGL---KSCMGLKIGUTVSISLEAKVRGQOE 409
FT	4 CYSTEINE-RICH TANDEM REPEATS.	Db	380 ELLENSESKREGVYISYRSECKNG--VKGTRGERGKCSNISGUQVERETSYTAHKCPKK 437
FT	REPEAT 466 515	Qy	410 -KEKSFТИKPGFKDSLIVTQYTFCDACQAKQEPBNRHCNNGTFCGVERCGPWLQ 468
FT	REPEAT 560 598	Db	438 GQAEISKIRPLGFNBEVEIWLQFICECDQDKGTPNSPCCHGNGTFCGAGRCNEGRIG 497
FT	DISULFID 28 464	Qy	469 SQCCESEEYRPSQCODE-CPSPREQPSQVPSQRGSCCLCGCQVCHSSDFGN --ITGKCCDD 525
FT	DISULFID 36 46	Db	498 KECECSTDEVNSEMDAYCRRENSEICNSNGDCIGCQVCKKKRDNPNPNEWVSKYCBDN 557
FT	DISULFID 39 76	Qy	526 FSCVYKJGENCSGHQOCGCLCQDSDWMTGYYCNCNTTRDTCSSNGLCISGRKGKCEGS 585
FT	DISULFID 49 65	Db	558 FNCDRSNGLGKGGKVKCRVCBEPNPGSAODCSDSTCMKNGCICRGICGDR 617
FT	DISULFID 206 212	Qy	586 CYCLOPGSYDTCBECPTOPDADCFKKEVCKFDRGALDENTCNRYCR---DELES 641
FT	DISULFID 260 300	Db	618 CKCTDPKFKQGPTCELCQTCVGVCTERHEKCVQCRFAQKGEQD--VCMEQCMHENLISLDS 675
FT	DISULFID 400 414	Qy	642 VKELDTG--KDAVNTCYKNEEDCVRPOYEESSGKSLILVVVEPCEPKGPDLVYLIS 699
FT	DISULFID 434 691	Db	676 REELPQPGQEAHTCHERKAEDCWYFVTSVSVSKNEVWVHVKEPECPGPDLIPIVAG 735
FT	DISULFID 462 489	Qy	700 VMGATLIGIAALLWKLITIHKRKEFKFEEBARAKWDTANNPJKEATSTFNITY 759
FT	DISULFID 477 525	Db	736 VVAGVLIGALLWKLIMIHDORREFAKEKMNAKWDGENPIVYKAVAVVNPYK 795
FT	DISULFID 486 525	Qy	760 RG 761
FT	DISULFID 491 500	Db	796 EG 797
FT	DISULFID 502 516	Qy	798 AA;
FT	DISULFID 531 536	Db	88303 MW; A1C5EA3711CF7C7 CRC64;
FT	DISULFID 533 568	OS	Mus musculus (Mouse);
FT	DISULFID 538 553	OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
FT	DISULFID 555 560	OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
FT	DISULFID 574 579	OC	NCBI_Taxid=10090;
FT	DISULFID 576 607	OC	[1] NCBITaxon=10090;
FT	DISULFID 581 590	RP	SEQUENCE FROM N. A.
FT	DISULFID 592 599	RC	STRAT=BA1B/C;
FT	DISULFID 613 618	RX	MEDLINE=82005707; PUBMED=3262537;
FT	DISULFID 615 661	RA	Tominaga S;
FT	DISULFID 620 630	RT	"Mutine mRNA for the beta-subunit of integrin is increased in
FT	DISULFID 633 636	RT	"Mutine mRNA for the beta-subunit of integrin is increased in
FT	DISULFID 640 649	RL	BALB/c-3T3 cells entering the G1 phase from the G0 state.";
FT	DISULFID 646 673	RN	[2] FEBS Lett. 238:315-319(1988).
FT	DISULFID 665 699	RP	SEQUENCE OF 2-798 FROM N. A.
FT	MOD_RES 783 783	RC	STRAT=BA1B/C;
FT	CARBOHYD 109 109	RC	MEDLINE=82235500; PubMed=2523953;
FT	CARBOHYD 131 131	RC	RE
FT	CARBOHYD 211 211	RC	RE
FT	CARBOHYD 223 223	RC	RE
FT	CARBOHYD 229 228	RC	RE
FT	CARBOHYD 268 362	RC	RE
FT	CARBOHYD 416 416	RC	RE
FT	CARBOHYD 481 520	RC	RE
FT	CARBOHYD 520 584	RC	RE
FT	CARBOHYD 584 669	RC	RE
FT	SEQUENCE 798 88303	OS	MEDLINE=82235500; PubMed=2523953;
Query	Match Score 44.0%; DB 1; Length 798; Best Local Similarity 44.8%; Pred. No. 2.4e-116; Matches 350; Conservative 127; Mismatches 276; Indels 29; Gaps 15;	Qy	1 GPMICTTRGVSSQOCLAVSPMCAWSD-EALPLGSP--RCDLKENLKDNCAPSSIEF 56
Qy	24 GGECLKANAKSGCCECQAGPNWGWCQVKYDQEGEPTASRCDLALKTKGKPCDION 83	Qy	57 PVSEARVLEDPRISLDSKGSG---DSQVTVQSPRQRIALRLRDDDSKNSIVQROVEDYVD 113
Db	84 PRQRKQKLKDIPITSKGKGERMDPANTQLRQPOQVFLVRLSPGEQPTNLFRRAEDYD 143	Qy	114 IYIIMDAYSYSMDLWSTONLGTQKTAQMRKTSNLRIGRAFDVDPVSPYMYISPEAL 173

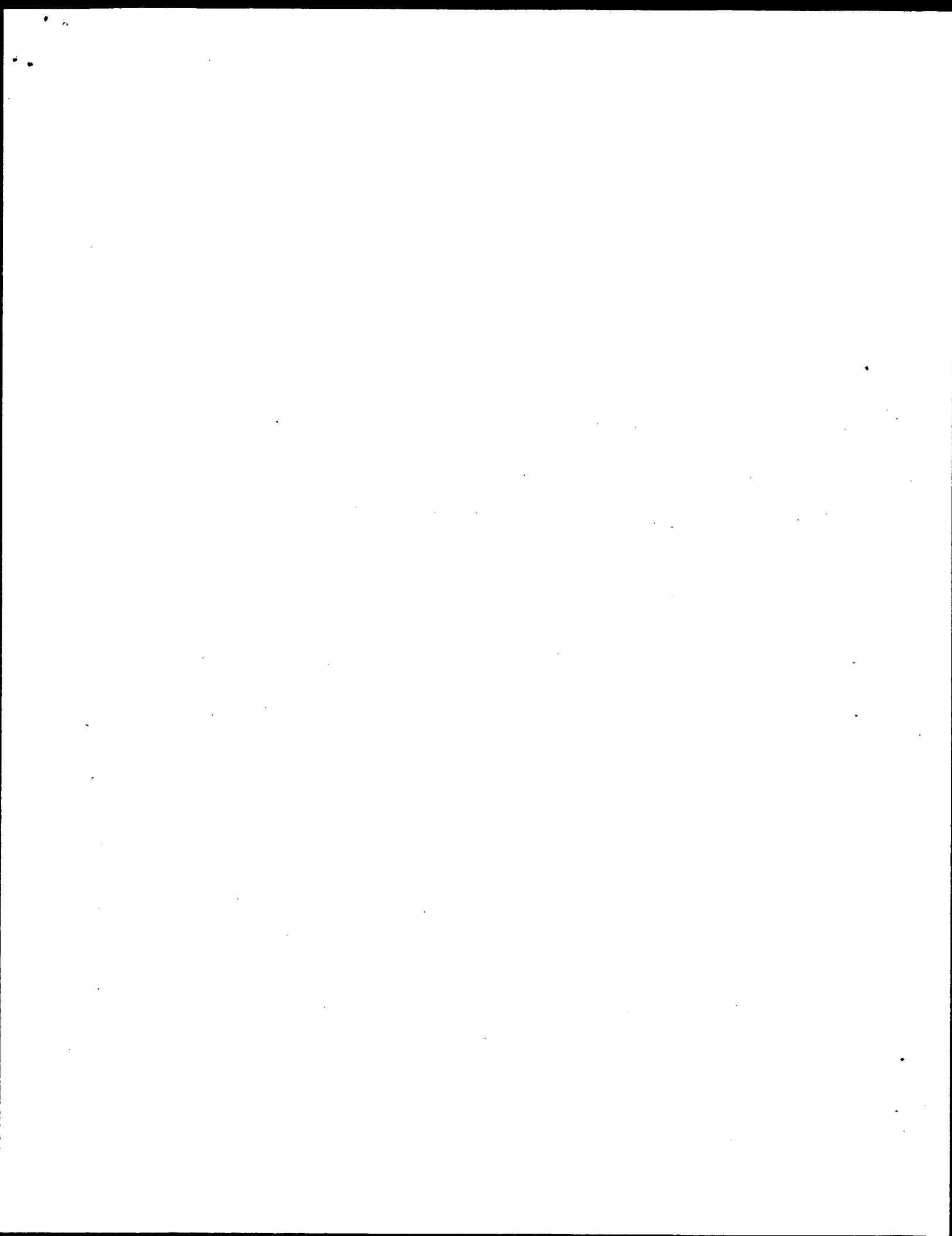
RA	Holers V.M., Ruff T.G., Parks D.L., McDonald J.A., Ballard L.L., Brown E.J.;	FT	DISULFID	27	464	BY SIMILARITY.
RT	"Molecular cloning of a murine fibronectin receptor and its expression during inflammation. Expression of vla-5 is increased in activated peritoneal macrophages in a manner discordant from major histocompatibility complex class II. ";	FT	DISULFID	35	45	BY SIMILARITY.
RT	J. EXP. Med. 169:1559-1605(1989).	FT	DISULFID	38	75	BY SIMILARITY.
CC	-1- FUNCTION: INTEGRINS ALPHA-1/BETA-1, ALPHA-2/BETA-1, ALPHA-10/BETA-1 AND ALPHA-11/BETA-1 ARE RECEPTORS FOR COLLAGEN. INTEGRINS ALPHA-1/BETA-1 AND ALPHA-2/BETA-2 RECOGNIZE THE PROLINE-HYDROXYLATED SEQUENCE G-F-P-G-E-R IN COLLAGEN. INTEGRINS ALPHA-2/BETA-1, ALPHA-3/BETA-1, ALPHA-4/BETA-1, ALPHA-1, ALPHA-8/BETA-1, ALPHA-10/BETA-1, ALPHA-11/BETA-1 AND ALPHA-V/BETA-1 ARE RECEPTORS FOR FIBRONECTIN. INTEGRIN ALPHA-4/BETA-1 RECOGNIZES ONE OR MORE DOMAINS WITHIN THE ALTERNATIVELY SPliced CS-1 AND CS-5 REGIONS OF FIBRONECTIN. INTEGRIN ALPHA-5/BETA-1 IS A RECEPTOR FOR FIBRINOGEN. INTEGRIN ALPHA-1/BETA-1, ALPHA-2/BETA-1, ALPHA-6/BETA-1 AND ALPHA-7/BETA-1 ARE RECEPTORS FOR LAMININ. INTEGRIN ALPHA-V/BETA-1 IS A RECEPTOR FOR VITROCORNEIN. INTEGRINS RECOGNIZE THE SEQUENCE R-G-D IN A WIDE ARRAY OF LIGANDS.	FT	DISULFID	401	415	BY SIMILARITY.
CC	-1- SUBUNIT: HETEROODIMER OF AN ALPHA AND A BETA SUBUNIT. BETA-1 ASSOCIATES WITH EITHER ALPHA-1, ALPHA-2, ALPHA-3, ALPHA-4, ALPHA-5, ALPHA-6, ALPHA-7, ALPHA-8, ALPHA-9, ALPHA-10, ALPHA-11 OR ALPHA-V.	FT	DISULFID	435	691	BY SIMILARITY.
CC	-1- SUBCELLULAR LOCATION: TYPE I membrane protein.	FT	DISULFID	462	465	BY SIMILARITY.
CC	-1- SIMILARITY: BELONGS TO THE INTEGRIN BETA CHAIN FAMILY.	FT	DISULFID	477	489	BY SIMILARITY.
CC	-1- SIMILARITY: CONTAINS 1 WW domain.	FT	DISULFID	500	525	BY SIMILARITY.
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement. (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).	FT	DISULFID	502	515	BY SIMILARITY.
CC	CC	FT	DISULFID	531	535	BY SIMILARITY.
CC	CC	FT	DISULFID	533	563	BY SIMILARITY.
CC	CC	FT	DISULFID	538	553	BY SIMILARITY.
CC	CC	FT	DISULFID	555	560	BY SIMILARITY.
CC	CC	FT	DISULFID	574	579	BY SIMILARITY.
CC	CC	FT	DISULFID	576	607	BY SIMILARITY.
CC	CC	FT	DISULFID	581	590	BY SIMILARITY.
CC	CC	FT	DISULFID	592	599	BY SIMILARITY.
CC	CC	FT	DISULFID	613	618	BY SIMILARITY.
CC	CC	FT	DISULFID	661	670	BY SIMILARITY.
CC	CC	FT	DISULFID	670	679	BY SIMILARITY.
CC	CC	FT	DISULFID	673	680	BY SIMILARITY.
CC	CC	FT	DISULFID	680	699	BY SIMILARITY.
CC	CC	FT	DISULFID	699	708	BY SIMILARITY.
CC	CC	FT	DISULFID	708	717	BY SIMILARITY.
CC	CC	FT	DISULFID	723	732	BY SIMILARITY.
CC	CC	FT	DISULFID	732	741	BY SIMILARITY.
CC	CC	FT	DISULFID	741	750	BY SIMILARITY.
CC	CC	FT	DISULFID	750	759	BY SIMILARITY.
CC	CC	FT	DISULFID	759	768	BY SIMILARITY.
CC	CC	FT	DISULFID	768	777	BY SIMILARITY.
CC	CC	FT	DISULFID	777	786	BY SIMILARITY.
CC	CC	FT	DISULFID	786	795	BY SIMILARITY.
CC	CC	FT	DISULFID	795	804	BY SIMILARITY.
CC	CC	FT	DISULFID	804	813	BY SIMILARITY.
CC	CC	FT	DISULFID	813	822	BY SIMILARITY.
CC	CC	FT	DISULFID	822	831	BY SIMILARITY.
CC	CC	FT	DISULFID	831	840	BY SIMILARITY.
CC	CC	FT	DISULFID	840	849	BY SIMILARITY.
CC	CC	FT	DISULFID	849	858	BY SIMILARITY.
CC	CC	FT	DISULFID	858	867	BY SIMILARITY.
CC	CC	FT	DISULFID	867	876	BY SIMILARITY.
CC	CC	FT	DISULFID	876	885	BY SIMILARITY.
CC	CC	FT	DISULFID	885	894	BY SIMILARITY.
CC	CC	FT	DISULFID	894	903	BY SIMILARITY.
CC	CC	FT	DISULFID	903	912	BY SIMILARITY.
CC	CC	FT	DISULFID	912	921	BY SIMILARITY.
CC	CC	FT	DISULFID	921	930	BY SIMILARITY.
CC	CC	FT	DISULFID	930	939	BY SIMILARITY.
CC	CC	FT	DISULFID	939	948	BY SIMILARITY.
CC	CC	FT	DISULFID	948	957	BY SIMILARITY.
CC	CC	FT	DISULFID	957	966	BY SIMILARITY.
CC	CC	FT	DISULFID	966	975	BY SIMILARITY.
CC	CC	FT	DISULFID	975	984	BY SIMILARITY.
CC	CC	FT	DISULFID	984	993	BY SIMILARITY.
CC	CC	FT	DISULFID	993	1002	BY SIMILARITY.
CC	CC	FT	DISULFID	1002	1011	BY SIMILARITY.
CC	CC	FT	DISULFID	1011	1020	BY SIMILARITY.
CC	CC	FT	DISULFID	1020	1029	BY SIMILARITY.
CC	CC	FT	DISULFID	1029	1038	BY SIMILARITY.
CC	CC	FT	DISULFID	1038	1047	BY SIMILARITY.
CC	CC	FT	DISULFID	1047	1056	BY SIMILARITY.
CC	CC	FT	DISULFID	1056	1065	BY SIMILARITY.
CC	CC	FT	DISULFID	1065	1074	BY SIMILARITY.
CC	CC	FT	DISULFID	1074	1083	BY SIMILARITY.
CC	CC	FT	DISULFID	1083	1092	BY SIMILARITY.
CC	CC	FT	DISULFID	1092	1101	BY SIMILARITY.
CC	CC	FT	DISULFID	1101	1110	BY SIMILARITY.
CC	CC	FT	DISULFID	1110	1119	BY SIMILARITY.
CC	CC	FT	DISULFID	1119	1128	BY SIMILARITY.
CC	CC	FT	DISULFID	1128	1137	BY SIMILARITY.
CC	CC	FT	DISULFID	1137	1146	BY SIMILARITY.
CC	CC	FT	DISULFID	1146	1155	BY SIMILARITY.
CC	CC	FT	DISULFID	1155	1164	BY SIMILARITY.
CC	CC	FT	DISULFID	1164	1173	BY SIMILARITY.
CC	CC	FT	DISULFID	1173	1182	BY SIMILARITY.
CC	CC	FT	DISULFID	1182	1191	BY SIMILARITY.
CC	CC	FT	DISULFID	1191	1200	BY SIMILARITY.
CC	CC	FT	DISULFID	1200	1209	BY SIMILARITY.
CC	CC	FT	DISULFID	1209	1218	BY SIMILARITY.
CC	CC	FT	DISULFID	1218	1227	BY SIMILARITY.
CC	CC	FT	DISULFID	1227	1236	BY SIMILARITY.
CC	CC	FT	DISULFID	1236	1245	BY SIMILARITY.
CC	CC	FT	DISULFID	1245	1254	BY SIMILARITY.
CC	CC	FT	DISULFID	1254	1263	BY SIMILARITY.
CC	CC	FT	DISULFID	1263	1272	BY SIMILARITY.
CC	CC	FT	DISULFID	1272	1281	BY SIMILARITY.
CC	CC	FT	DISULFID	1281	1290	BY SIMILARITY.
CC	CC	FT	DISULFID	1290	1299	BY SIMILARITY.
CC	CC	FT	DISULFID	1299	1308	BY SIMILARITY.
CC	CC	FT	DISULFID	1308	1317	BY SIMILARITY.
CC	CC	FT	DISULFID	1317	1326	BY SIMILARITY.
CC	CC	FT	DISULFID	1326	1335	BY SIMILARITY.
CC	CC	FT	DISULFID	1335	1344	BY SIMILARITY.
CC	CC	FT	DISULFID	1344	1353	BY SIMILARITY.
CC	CC	FT	DISULFID	1353	1362	BY SIMILARITY.
CC	CC	FT	DISULFID	1362	1371	BY SIMILARITY.
CC	CC	FT	DISULFID	1371	1380	BY SIMILARITY.
CC	CC	FT	DISULFID	1380	1389	BY SIMILARITY.
CC	CC	FT	DISULFID	1389	1398	BY SIMILARITY.
CC	CC	FT	DISULFID	1398	1407	BY SIMILARITY.
CC	CC	FT	DISULFID	1407	1416	BY SIMILARITY.
CC	CC	FT	DISULFID	1416	1425	BY SIMILARITY.
CC	CC	FT	DISULFID	1425	1434	BY SIMILARITY.
CC	CC	FT	DISULFID	1434	1443	BY SIMILARITY.
CC	CC	FT	DISULFID	1443	1452	BY SIMILARITY.
CC	CC	FT	DISULFID	1452	1461	BY SIMILARITY.
CC	CC	FT	DISULFID	1461	1470	BY SIMILARITY.
CC	CC	FT	DISULFID	1470	1479	BY SIMILARITY.
CC	CC	FT	DISULFID	1479	1488	BY SIMILARITY.
CC	CC	FT	DISULFID	1488	1497	BY SIMILARITY.
CC	CC	FT	DISULFID	1497	1506	BY SIMILARITY.
CC	CC	FT	DISULFID	1506	1515	BY SIMILARITY.
CC	CC	FT	DISULFID	1515	1524	BY SIMILARITY.
CC	CC	FT	DISULFID	1524	1533	BY SIMILARITY.
CC	CC	FT	DISULFID	1533	1542	BY SIMILARITY.
CC	CC	FT	DISULFID	1542	1551	BY SIMILARITY.
CC	CC	FT	DISULFID	1551	1560	BY SIMILARITY.
CC	CC	FT	DISULFID	1560	1569	BY SIMILARITY.
CC	CC	FT	DISULFID	1569	1578	BY SIMILARITY.
CC	CC	FT	DISULFID	1578	1587	BY SIMILARITY.
CC	CC	FT	DISULFID	1587	1596	BY SIMILARITY.
CC	CC	FT	DISULFID	1596	1605	BY SIMILARITY.
CC	CC	FT	DISULFID	1605	1614	BY SIMILARITY.
CC	CC	FT	DISULFID	1614	1623	BY SIMILARITY.
CC	CC	FT	DISULFID	1623	1632	BY SIMILARITY.
CC	CC	FT	DISULFID	1632	1641	BY SIMILARITY.
CC	CC	FT	DISULFID	1641	1650	BY SIMILARITY.
CC	CC	FT	DISULFID	1650	1659	BY SIMILARITY.
CC	CC	FT	DISULFID	1659	1668	BY SIMILARITY.
CC	CC	FT	DISULFID	1668	1677	BY SIMILARITY.
CC	CC	FT	DISULFID	1677	1686	BY SIMILARITY.
CC	CC	FT	DISULFID	1686	1695	BY SIMILARITY.
CC	CC	FT	DISULFID	1695	1704	BY SIMILARITY.
CC	CC	FT	DISULFID	1704	1713	BY SIMILARITY.
CC	CC	FT	DISULFID	1713	1722	BY SIMILARITY.
CC	CC	FT	DISULFID	1722	1731	BY SIMILARITY.
CC	CC	FT	DISULFID	1731	1740	BY SIMILARITY.
CC	CC	FT	DISULFID	1740	1749	BY SIMILARITY.
CC	CC	FT	DISULFID	1749	1758	BY SIMILARITY.
CC	CC	FT	DISULFID	1758	1767	BY SIMILARITY.
CC	CC	FT	DISULFID	1767	1776	BY SIMILARITY.
CC	CC	FT	DISULFID	1776	1785	BY SIMILARITY.
CC	CC	FT	DISULFID	1785	1794	BY SIMILARITY.
CC	CC	FT	DISULFID	1794	1803	BY SIMILARITY.
CC	CC	FT	DISULFID	1803	1812	BY SIMILARITY.
CC	CC	FT	DISULFID	1812	1821	BY SIMILARITY.
CC	CC	FT	DISULFID	1821	1830	BY SIMILARITY.
CC	CC	FT	DISULFID	1830	1839	BY SIMILARITY.
CC	CC	FT	DISULFID	1839	1848	BY SIMILARITY.
CC	CC	FT	DISULFID	1848	1857	BY SIMILARITY.
CC	CC	FT	DISULFID	1857	1866	BY SIMILARITY.
CC	CC	FT	DISULFID	1866	1875	BY SIMILARITY.
CC	CC	FT	DISULFID	1875	1884	BY SIMILARITY.
CC	CC	FT	DISULFID	1884	1893	BY SIMILARITY.
CC	CC	FT	DISULFID	1893	1902	BY SIMILARITY.
CC	CC	FT	DISULFID	1902	1911	BY SIMILARITY.
CC	CC	FT	DISULFID	1911	1920	BY SIMILARITY.
CC	CC	FT	DISULFID	1920	1929	BY SIMILARITY.
CC	CC	FT	DISULFID	1929	1938	BY SIMILARITY.
CC	CC	FT	DISULFID	1938	1947	BY SIMILARITY.
CC	CC	FT	DISULFID	1947	1956	BY SIMILARITY.
CC	CC	FT	DISULFID	1956	1965	BY SIMILARITY.
CC	CC	FT	DISULFID	1965	1974	BY SIMILARITY.
CC	CC	FT	DISULFID	1974	1983	BY SIMILARITY.
CC	CC	FT	DISULFID	1983	1992	BY SIMILARITY.
CC	CC	FT	DISULFID	1992	2001	BY SIMILARITY.
CC	CC	FT	DISULFID	2001	2010	BY SIMILARITY.
CC	CC	FT	DISULFID	2010	2019	BY SIMILARITY.
CC	CC	FT	DISULFID	2019	2028	BY SIMILARITY.
CC	CC	FT	DISULFID	2028	2037	BY SIMILARITY.
CC	CC	FT	DISULFID	2037	2046	BY SIMILARITY.
CC	CC	FT	DISULFID	2046	2055	BY SIMILARITY.
CC	CC	FT	DISULFID	2055	2064	BY SIMILARITY.
CC	CC	FT	DISULFID	2064	2073	BY SIMILARITY.
CC	CC	FT	DISULFID	2073	2082	BY SIMILARITY.
CC	CC	FT	DISULFID	2082	2091	BY SIMILARITY.
CC	CC	FT	DISULFID	2091	21	

CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; U12309; AA86669.1; -
 DR InterPro; IPR00561; EGF-like.
 DR InterPro; IPR002369; Integrin_B.
 DR InterPro; IPR00169; Integrin_beta_C.
 DR InterPro; IPR03659; PSI.
 DR InterPro; IPR002035; wWA.
 DR Pfam; PF00362; integrin_B. 1.
 DR PRINS; PRO180; INTEGRIN_B.
 DR PROBON; P001811; Integrin_B. 1.
 DR SMART; SM00187; EGF-like. 1.
 DR SMART; SM00423; PSI. 1.
 DR PROSITE; PS00343; INTEGRIN_BETA. 3.
 DR PROSITE; PS00022; EGF1; UNKNOWN. 2.
 KW Integrin; Cell adhesion; Receptor; Transmembrane; Glycoprotein;
 KW Repeat; Signal; Phosphorylation.
 FT SIGNAL 1 20 BY SIMILARITY.
 FT CHAIN 21 798 INTEGRIN_BETA.1.
 FT DOMAIN 21 729 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 730 752 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 140 378 wWA-LIKE.
 FT DOMAIN 467 636 4 CYSTEINE-RICH TANDEM REPEATS.
 FT REPEAT 517 560 II.
 FT REPEAT 561 599 III.
 FT REPEAT 600 636 IV.
 FT DISUFLID 27 465 BY SIMILARITY.
 FT DISUFLID 35 45 BY SIMILARITY.
 FT DISUFLID 38 75 BY SIMILARITY.
 FT DISUFLID 48 64 BY SIMILARITY.
 FT DISUFLID 207 213 BY SIMILARITY.
 FT DISUFLID 261 301 BY SIMILARITY.
 FT DISUFLID 401 415 BY SIMILARITY.
 FT DISUFLID 435 692 BY SIMILARITY.
 FT DISUFLID 463 467 BY SIMILARITY.
 FT DISUFLID 478 490 BY SIMILARITY.
 FT DISUFLID 487 526 BY SIMILARITY.
 FT DISUFLID 492 501 BY SIMILARITY.
 FT DISUFLID 503 517 BY SIMILARITY.
 FT DISUFLID 532 537 BY SIMILARITY.
 FT DISUFLID 534 569 BY SIMILARITY.
 FT DISUFLID 539 554 BY SIMILARITY.
 FT DISUFLID 556 561 BY SIMILARITY.
 FT DISUFLID 575 580 BY SIMILARITY.
 FT DISUFLID 577 608 BY SIMILARITY.
 FT DISUFLID 582 591 BY SIMILARITY.
 FT DISUFLID 593 600 BY SIMILARITY.
 FT DISUFLID 614 619 BY SIMILARITY.
 FT DISUFLID 616 662 BY SIMILARITY.
 FT DISUFLID 621 631 BY SIMILARITY.
 FT DISUFLID 634 637 BY SIMILARITY.
 FT DISUFLID 641 650 BY SIMILARITY.
 FT DISUFLID 647 724 BY SIMILARITY.
 FT DISUFLID 666 700 BY SIMILARITY.
 FT MOD_RES 784 784 PHOSPHORYLATION (BY SIMILARITY).
 FT CARBOHYD 50 50 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 94 94 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 97 97 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 212 212 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 269 269 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 363 363 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 406 406 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 417 417 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 482 482 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 521 521 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 585 585 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 670 670 N-LINKED (GLCNAC. . .) (POTENTIAL).

Query Match 43.4%; Score 1801.5; DB 1; Length 799;
 Best Local Similarity 44.2%; Preced. No. 1.2e-114; Mismatches 271; Index 29; Gaps 14;
 Matches 345; Conservative 136; MisMatches 271; Index 29; Gaps 14;
 QY 3 NICCTTRGVSSCOCLAVSPMCACSD----EALPESGSPRCDLKENLLKNCAPESIEFP 57
 Db 25 NRCLAKANAKSCGECIQAGPNCGCWCTINTFLOGMPT-SARCDDELAKKKGCHPSDIEP 83
 QY 58 VSEARVEDRPLSKGDSS-----VTOVSQFALRIPDRDKNSFSTQYRVEDPVY 112
 Db 84 RGSQTIRKRNKNVNRSKGMAERLPRDITQIQPQQQLKLRSGEPOQKFTKRAEDYPI 143
 QY 113 DIYVLMOLSYSKADDLNSIQNIGKLTATOMKLTSLRIGFAFDVKPSVPMISP 172
 Db 144 DAXYLMOLSYSKADDLNSIQNIGKLTATOMKLTSLRIGFAFDVKPSVPMISP 202
 QY 173 LENPCYDMMTCTOLPMQKXHVILTQVTRNEEYKQSVRNRAPEGGFDATMQAVC 232
 Db 203 LRNPCT-TSEQNQTSPPSYKVLNSLDPGEFFENLYQDRTSGNLDSPPEGFDATMQAVC 261
 QY 233 DEKFTGWNDAHSLVLTDTAKTHIALDGLLAGIVOPNDGQCHVGSDNHYASASTMDYPSL 292
 Db 262 GSLIGWRN-VTRLLVFTSDAGFHAGFQKGKLGIVPLNDGQCHL-ENNYVITMSHYDYSI 319
 QY 293 GLMTTEKISQSKNQNLTFATWENVNLQNSYSELIPCTVGLSMSSNVIOLIVAYGKIR 352
 Db 320 AHLVQKISSENNIQTFAVTEFQPVKLNKLPISAVGTLISGNNSNVIOLITDAYSN 379
 QY 353 SKVELEVRDLEBELSFSNATCLN--NEVTPGLKSCMGLKIGDTVSSTEAKVRGCP QE 409
 Db 380 SEVILENSKLPDGVTINYKSYKSYKNGVNGTENGKRCNSNTSIGDEVOFETISATANKCPNE 439
 QY 410 KEKFTIIPVQFKDSLIVQVIFDCDCAQCAQAEPAEPNSHRCNGNGTFCGGVRCRGPGWGS 469
 Db 440 SENOLKNUPLGFTEEVVLQFICKNQSGHIPASPKCHEGNQTEFGACRNCNEGRGVR 499
 QY 470 QCECSERBDYRPSQODE-CSPREGQPVCSQREGCICGQCVCHSSDFGK--ITGKQCECDF 526
 Db 500 HCECSTDDEVNSEDMDAYCRKENSSETCSNNNGCVCQGCVCRRKRENTNEYISGACECDNF 559
 QY 527 SCVRYKGEMCSGHGCGCSDCSDWTSYCNCTTRTDMCSNSGLLCSGRKCECCSC 586
 Db 560 NCDRSNGLICCGNGVCRVRCYCYPNTESACDCSLSLTVPCVATNGQICNGRGCECCAC 619
 QY 587 VCIQPSYQDCEKCPCTPDACTFFKKECVCCKKFDRGALDENTCNRCR--DEFSV 642
 Db 620 KCTDPKFGPCTCOTCGLVCAEHKEVQCAFRAENKGKD--TCAQCSHFNLTQVSR 677
 QY 643 KELKDGTGK--DAVNCTYKNEEDCVVRFQTYEDSSGKSLXVVERPECKRGPDILVLSV 700
 Db 678 EKLQPQVQDVPTVHCKEDIDCWYFTSVNSKGAEAHVWETPDCTGPDTIPIVAGV 737
 QY 701 MGAIIILIGLALLWKLITTHDRKEFAFEERAKNDTANNPYKEATSTFTNTYR 760
 Db 738 VAGIVLVLGLALLLWKLMLIHDREFAKEFEKRNKAQWTGNGENPIYSAVTVVNPKY 797
 QY 761 G 761
 Db 798 G 798

Search completed: May 19, 2002, 12:27:36
 Job time: 6470 sec



GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on:

May 19, 2002, 10:38:58 ; Search time 96.49 Seconds

(without alignments)
1366.174 Million cell updates/sec

Title:

US-09-673-302a-1

Perfect score:

4154 1 GENDCT"RGVSSCQCLAVS.....NNPLYKEATSTFTNITYRGT 762

Sequence:

BLOSUM62

Scoring table: GapOp 10.0 , GapExt 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 10%

Listing first 45 summaries

Database : SPTREMBL_19:*

- 1: sp_archea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_orangelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rabbit:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriapl:*
- 17: sp_archeap:*

Pred. No. 18 is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

RESULT	1
ID	097702
PRELIMINARY:	
AC	097702;
DT	01-MAY-1999 (TREMBrel. 10, Created)
DT	01-MAY-1999 (TREMBrel. 10, Last sequence update)
DT	01-JDC-2001 (TREMBrel. 19, Last annotation update)
DE	PLATELET GLYCOPROTEIN IIIA.
GN	GP11A.
OS	Canis familiaris (Dog).
OC	Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX	NCBI-TaxID=9615;
RN	[1]
RP	SEQUENCE FROM N.A.
RX	MEDLINE=9541003; PUBMED=10482317;
RA	Lipscomb D.L., Bourne C., Boudreault M.K.;
RT	"DNA sequence of the canine platelet beta3 gene from cDNA: comparison of canine and mouse beta3 to segments that encode alloantigenic sites and functional domains of beta3 in human beings";
RT	RJ. Lab. Clin. Med. 134:313-321(1999).
CC	-1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN ('BY SIMILARITY').
CC	-1- PPM: THE CYSTEINE RESIDUES ARE INVOLVED IN INTRACHAIN DISULFIDE BONDS (BY SIMILARITY).
CC	-1- SIMILARITY: BELONGS TO THE INTEGRIN BETA CHAIN FAMILY.
DR	EMBL: AF116270; AAD13680.1; -.
DR	INTERPRO: IPR002086; Alpha_beta_dehydr.
DR	INTERPRO: IPR00561; EGF-like.
DR	INTERPRO: IPR02369; Integrin_B.
DR	INTERPRO: IPR00169; Integrin_beta_C.
DR	INTERPRO: IPR03659; S1.
DR	INTERPRO: IPR02035; vWFA.
DR	PFAM: PF00362; integrin_B.1.
DR	PRINTS: PRO1185; INTEGRIN_B.
DR	PRODOM: PDO01811; Integrin_B.1.
DR	SMART: SM00001; EGF_like.1.
DR	SMART: SM0017; INB.1.
DR	SMART: SM00423; PSI.1.
DR	SMART: SM00327; VWA.1.

Result No.	Score	Query	Match length	DB ID	Description
1	4016	96.7	784	6 097702	097702 canis famili
2	4013	96.6	784	6 Q9TQ07	Q9TQ07 canis famili
3	3996	96.2	788	6 Q9TQ03	Q9TQ03 oryctolagus
4	3989	93.8	784	6 Q95JH1	Q95JH1 sus scrofa
5	3884	93.7	784	6 Q9TQ05	Q9TQ05 sus scrofa
6	3606	86.8	723	11 Q9QW16	Q9QW16 ractus sp.
7	3484	83.9	781	13 Q9Z071	Q9Z071 gallus gall
8	3472	83.6	680	11 Q9QW15	Q9QW15 mus sp. bet
9	3260	78.5	788	13 Q07012	Q07012 xenopus lae
10	2315	55.7	791	6 Q9K499	Q9K499 bos taurus
11	1833	44.1	798	6 Q9GJF0	Q9GJF0 sus scrofa
12	1800	43.3	807	13 Q9IA01	Q9IA01 ictalurus p
13	1658	39.9	509644		096444 bioluminesc
14	1609	38.7	837	5 Q9NA57	Q9NA57 anopheles g
15	1588	38.2	783	5 Q92163	Q92163 strongyloice
16	1584	38.1	799	5 Q95P95	Q95P95 crassostrea

DR	PROSITE; PS00070; ALDEHYDE_DEHYDR_CYS; UNKNOWN_1.	DE GLYCOPROTEIN GPIIIA.
DR	PROSITE; PS00022; EGF_1; UNKNOWN_2.	GN GPIIIA.
DR	PROSITE; PS01186; EGF_2; 1.	OS Canis familiaris (Dog).
KW	Cell adhesion; Cytoskeleton; EGF-like domain; Extracellular matrix; Glycoprotein; Integrin; Repeat; Transmembrane.	OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
SEQ	SEQUENCE 784 AA; 86388 MW; 6B2A6D34916EA260 CRC64;	OX NCB - ToxID=615;
	SEQUENCE FROM N.A.	RN [1]
	RP TISSUE=BLOOD;	RP
	RA Tissue=R;	RA
	RT Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.	RT
	CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (BY SIMILARITY).	CC
	CC -!- PTM: THE CYSTEINE RESIDUES ARE INVOLVED IN INTRACHAIN DISULFIDE BONDS (BY SIMILARITY).	CC
	CC -!- SIMILARITY: BELONGS TO THE INTEGRIN BETA CHAIN FAMILY.	CC
	DR EMBL; AF170525; AAD49373_1; -.	DR
	DR InterPro; IPR002086; Integrin_beta.	DR
	DR InterPro; IPR000561; EGF-like.	DR
	DR InterPro; IPR002369; Integrin_B.	DR
	DR InterPro; IPR001169; Integrin_beta_C.	DR
	DR InterPro; IPR013659; PSI.	DR
	DR InterPro; IPR002035; wWFA.	DR
	DR InterPro; IPR00162; Integrin_B.	DR
	DR Prints; PRO1186; INTEGRIN_B.	DR
	DR ProDom; PD00181; Integrin_B.	DR
	DR SMART; SM00001; EGF_like.	DR
	DR SMART; SM00187; INB.	DR
	DR SMART; SM00423; PSI.	DR
	DR SMART; SM00327; WWA.	DR
	DR PROSITE; PS00070; ALDEHYDE_DEHYDR_CYS; UNKNOWN_1.	DR
	DR PROSITE; PS00022; EGF_1; UNKNOWN_2.	DR
	DR PROSITE; PS01186; EGF_2; 1.	DR
	DR PROSITE; PS00243; INTEGRIN_BETA; 3.	DR
	DR Cell adhesion; Cytoskeleton; EGF-like domain; Extracellular matrix; Glycoprotein; Integrin; Repeat; Transmembrane.	DR
	SEQ SEQUENCE 784 AA; 86415 MW; 9D507F827628790A CRC64;	DR
	Query Match 96.7%; Score 4016; DB 6; Length 784;	Query Match 96.6%; Score 4013; DB 6; Length 784;
	Best Local Similarity 96.3%; Pred. No. 0; Mismatches 14; Indels 0; Gaps 0;	Best Local Similarity 96.2%; Pred. No. 0; Mismatches 15; Indels 0; Gaps 0;
	Matches 731; Conservative 95%; MisMatch 14; Insertions 0; Gaps 0;	Matches 731; Conservative 95%; MisMatch 15; Insertions 0; Gaps 0;
QY	3 NICTRGVSSCQCLAVSPMCAWCSDAELPLGSPRCDLKENLKDNCAPESIEFPVSEAR 62	QY 3 NICTRGVSSCQCLAVSPMCAWCSDAELPLGSPRCDLKENLKDNCAPESIEFPVSEAR 62
Db	25 NICTRGVHSQOCCLAVSPCWCSDAELPLGSPRCDLKENLKDNCAPESIEFPVSEVR 84	Db 25 NICTRGVHSQOCCLAVSPCWCSDAELPLGSPRCDLKENLKDNCAPESIEFPVSEVR 84
QY	63 VLEDRPLSDKGSGDSSQVQSPORIALRLRPPDSKNSIIVQROVEDYPVDIYLMDSY 122	QY 63 VLEDRPLSDKGSGDSSQVQSPORIALRLRPPDSKNSIIVQROVEDYPVDIYLMDSY 122
Db	85 ILEARPLSNKGSGDSSQVQSPORIALRLRPPDSKNSIIVQROVEDYPVDIYLMDSY 144	Db 85 ILEARPLSNKGSGDSSQVQSPORIALRLRPPDSKNSIIVQROVEDYPVDIYLMDSY 144
QY	123 SMKDDLWISONLGLTRKLTASQMKHLTSNRLRIGFGAFVDPVSKYPMYISPEALINPCYMKT 182	QY 123 SMKDDLWISONLGLTRKLTASQMKHLTSNRLRIGFGAFVDPVSKYPMYISPEALINPCYMKT 182
Db	145 SMKDDLWISONLGLTRKLTASQMKHLTSNRLRIGFGAFVDPVSKYPMYISPEALINPCYMKT 204	Db 145 SMKDDLWISONLGLTRKLTASQMKHLTSNRLRIGFGAFVDPVSKYPMYISPEALINPCYMKT 204
QY	183 TCLPMFGYKHVLTLTDQVTRFNEEVKKQSRSNRDAPEGGFADIMQATVCDKEIGWRDA 242	QY 183 TCLPMFGYKHVLTLTDQVTRFNEEVKKQSRSNRDAPEGGFADIMQATVCDKEIGWRDA 242
Db	205 TCLPMFGYKHVLTLTDQVTRFNEEVKKQSRSNRDAPEGGFADIMQATVCDKEIGWRDA 264	Db 205 TCLPMFGYKHVLTLTDQVTRFNEEVKKQSRSNRDAPEGGFADIMQATVCDKEIGWRDA 264
QY	243 SHLIVFTTDKTHALDGRLAGITQVQPNQDGCHGSDNHYKASSTMDYPSLGLMTEKUSQ 302	QY 243 SHLIVFTTDKTHALDGRLAGITQVQPNQDGCHGSDNHYKASSTMDYPSLGLMTEKUSQ 302
Db	265 SHLIVFTTDKTHALDGRLAGITQVQPNQDGCHGSDNHYKASSTMDYPSLGLMTEKUSQ 324	Db 265 SHLIVFTTDKTHALDGRLAGITQVQPNQDGCHGSDNHYKASSTMDYPSLGLMTEKUSQ 324
QY	303 NINLIFAVTENVNLVLYQNSSELIPGTTVGLSMOSSNNYQLOLTYDAYGRSKYEVLEVDL 362	QY 303 NINLIFAVTENVNLVLYQNSSELIPGTTVGLSMOSSNNYQLOLTYDAYGRSKYEVLEVDL 362
Db	325 NINLIFAVTENVNLVLYQNSSELIPGTTVGLSMOSSNNYQLOLTYDAYGRSKYEVLEVDL 384	Db 325 NINLIFAVTENVNLVLYQNSSELIPGTTVGLSMOSSNNYQLOLTYDAYGRSKYEVLEVDL 384
RESULT	2	
Q9TUN7	PRELIMINARY; PRT; 784 AA.	Q9TUN7 PRELIMINARY; PRT; 784 AA.
ID	Q9TUN7; 01-MAY-2000 (Tremblrel. 13, Created)	ID Q9TUN7; 01-MAY-2000 (Tremblrel. 13, Last sequence update)
AC	385 PEELSISNATCLNNENVIPGLKSCVGLKGIGDVFYPSFIRAKVRCGPQEKEKSFITKPVGFK	AC 385 PEELSISNATCLNNENVIPGLKSCVGLKGIGDVFYPSFIRAKVRCGPQEKEKSFITKPVGFK
DT	01-DEC-2001 (Tremblrel. 19, Last annotation update)	DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)

RESULT 3

Q9TUN3 PRELIMINARY; PRT: 788 AA.

ID Q9TUN3; 01-MAY-2000 (TREMBrel. 13, Created)

DT 01-DEC-2001 (TREMBrel. 19, Last annotation update)

DE GLYCOPROTEIN GPIIIA.

OS Oryctolagus cuniculus (Rabbit)

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.

OX NCBI_TaxID:9986;

RN [1] SEQUENCE FROM N.A.

RC TISSUE=PLATELET;

RA Tao J., Parilla R.;

RT Molecular cloning of rabbit platelet GPIIIA cDNA.;

RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.

CC -I SUBCELLULAR LOCALIZATION: TYPE I MEMBRANE PROTEIN (BY SIMILARITY).

CC -I PTM: THE CYSTEINE RESIDUES ARE INVOLVED IN INTRACHAIN DISULFIDE BONDS (BY SIMILARITY).

CC -I SIMILARITY: BELONGS TO THE INTEGRIN BETA CHAIN FAMILY.

DR EMBL; AF170529; AAD51955.1; -.

DR InterPro; IPR002086; Aldehyde_dehydr.

DR InterPro; IPR00561; EGF-like.

DR InterPro; IPR002369; Integrin_B.

DR InterPro; IPR003659; PSI.

DR InterPro; IPR002035; VWF.

DR InterPro; IPR00169; Integrin_beta_C.

DR InterPro; IPR003659; PSI.

DR InterPro; IPR002035; VWF.

DR PRINTS; PRO0162; Integrin_B. 1.

DR PRINTS; PRO0181; Integrin_B. 1.

DR SMART; SM00187; INB. 1.

DR SMART; SM00423; PSI. 1.

DR SMART; SM00327; VWA. 1.

DR PROSITE; PS00070; ALBHYDE_DEHYDR_CYS; UNKNOWN_1.

DR PROSITE; PS00022; EGF_1; UNKNOWN_2.

DR PROSITE; PS0186; EGF_2; 1.

DR PROSITE; PS00243; INTEGRIN_BETA; 3

KW Cell adhesion; Cytoskeleton; EGF-like domain; Extracellular matrix; Glycoprotein; Integrin; Repeat; Transmembrane.

KW SEQUENCE 788 AA; 87049 MW; 7AD9CDA7301D78D CRC64;

RESULT 4

Q95JH1 PRELIMINARY; PRT: 784 AA.

ID Q95JH1; 01-DEC-2001 (TREMBrel. 19, Created)

DT 01-DEC-2001 (TREMBrel. 19, Last annotation update)

DE GLYCOPROTEIN GPIIIA.

GN CD61.

OS Sus scrofa (P19).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.

OX NCBI_TaxID:9823;

RN [1] SEQUENCE FROM N.A.

RC TISSUE=AORTIC SMOOTH MUSCLE;

RA Jimenez-Marin A.M., Garrido J.J., Llanes D., Barbancho M.J.;

RT "Characterization of the porcine CD61 (GPIIIa) gene.";

RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF282890; ANK69529.1; -.

SQ SEQUENCE 784 AA; 86365 MW; 1AE73665IEF33B8 CRC64;

Query Match 96.2% score 3996; DB 6; Length 788; Best Local Similarity 95.3%; Pred. No. 0; Matches 726; Conservative 21; Mismatches 15; Indels 0; Gaps 0;

Q9 1 GPNICTRGVSSCQCLAVSPMCAMSDEALPLGSPRCOLKENLKDNCAPESIEFPVSE 60

RESULT		5	RP	SEQUENCE FROM N.A.
ID	Q9TUN5	PRELIMINARY;	RC	TISSUE=PLATELET;
AC	Q9TUN5;	PRT;	RA	Tao J., Parilla R.;
DT	01-MAY-2000 (TREMBLrel. 13, Created)		RT	"Molecular cloning of pig platelet GPIIa cDNA."
DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)		RL	Submitted (TUL-1999) to the EMBL/GenBank/DDBJ databases.
DT	01-DEC-2001 (TREMBLrel. 19, Last annotation update)		CC	-1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (BY SIMILARITY).
DE	GLYCOPROTEIN IIIA.		CC	BONDS (BY SIMILARITY).
GN	GPITTA.		CC	-1- SIMILARITY: BELONGS TO THE INTEGRIN BETA CHAIN FAMILY.
OS	Sus scrofa (Pig).		DR	EMBL; AR170527; AAD51535.1; -.
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		DR	InterPro; IPR02086; Aldehyde_dehydr.
OC	Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.		DR	InterPro; IPR00561; EGF-like.
OX	NCBL_TAXID=923;		DR	InterPro; IPR03006; Ig_MHC.
RN	{1}		DR	InterPro; IPR02369; Integrin_B.
Query Match		93.8%	Score	3898; DB 6; Length 784;
Best Local Similarity		92.8%	Pred. No.	0; Mismatches
Matches	707; Conservative	35;	Indels	0; Gaps
Qy	1	GPNCITRGVSSCQCLAVSPMCAMCSDAELPLGSPRCDLKENLKDCAPESTEFPSE	60	
Db	23	GPNCITRGVSSCQCLAVSPMCAMCSDAELPLGSPRCDLKENLKDCAPESTEFPSE	82	
Qy	61	ARVLEDPLSKGSQDSQVQSPORIALRLRDPDSKNISIQVROVEDYPVDIYMLD	120	
Db	83	ARILEAPLSPSKGSQDSQVQSPORIALRLRDPDSKNISIQVROVEDYPVDIYMLD	142	
Qy	181	KTTCIPLMFGYKHVLTLDQVTRFNEEVKKQSVSRNRDAPEGGDAIMQATVCDEKIGRN	240	
Db	203	KTTCIPLMFGYKHVLTLDQVTRFNEEVKKQSVSRNRDAPEGGDAIMQATVCDEKIGRN	262	
Qy	241	DASHLVLVFTDAKTHIALDGLLAGTVQPNQGCHGSDNHYASASTMYPQLGIMTEKL	300	
Db	263	DASHLVLVFTDAKTHIALDGLLAGTVQPNQGCHGSDNHYASASTMYPQLGIMTEKL	322	
Qy	301	QKNINLIFAVTENVLYQNSELIPGTVQPNQGCHGSDNHYASASTMYPQLGIMTEKL	360	
Db	323	QKNINLIFAVTENVLYQNSELIPGTVQPNQGCHGSDNHYASASTMYPQLGIMTEKL	382	
Qy	361	DLPEELSLSFNATCLNEVTPGKSMGKIGDVSISIEAKYRGCPQEKEKSTIPIVG	420	
Db	383	DLPEELSLSFNATCLNEVTPGKSMGKIGDVSISIEAKYRGCPQEKEKSTIPIVG	442	
Qy	421	FKDSLTVQVTFNCDCCQAOAEPDSPRCDNGTFECGVRCGPGWLGSQCESEEDYRP	480	
Db	443	FKDSLTVQVTFNCDCCQAOAEPDSPRCDNGTFECGVRCGPGWLGSQCESEEDYRP	502	
Qy	481	SQDDECSPREPGQPVCSORGCLCQCVCHSSDGFKITGYCECDDFSCVYKGCMSHG	540	
Db	503	SQDDECSPREPGQPVCSORGCLCQCVCHSSDGFKITGYCECDDFSCVYKGCMSHG	562	
Qy	541	QCSGCDLCLSDWYGYYCCTTRDTCMSANGLICSGRGKCEGSCGVCIQPSYGDCEK	600	
Db	563	QCSGCDLCLSDWYGFYCCTTRDTCMSANGLICSGRGKCEGSCGVCIQPSYGDCEK	622	
Qy	601	CPTCPDACTFKECVECKKFDRALHDENTCNYCDTELETSYKELDKGKAVNCTYNE	660	
Db	623	CPTCDACITFKKGCVCKKFDRALHDENTCNYCDTELETSYKELDKGKAVNCTYNE	682	
Qy	661	DDCVYRFOYEDSSGKSLIVVVERPECKGPDILVLLSVMGAILLIGLALLIKLIT	720	
Db	683	DDCVYRFOYEDSSGKSLIVVVERPECKGPDILVLLSVMGAILLIGLALLIKLIT	742	
Qy	721	IHDKEFAKEFEERAKWDTANNPPLYKEATSIFTNTYRGT	762	
Db	743	IHDKEFAKEFEERAKWDTANNPPLYKEATSIFTNTYRGT	784	
Query Match		93.7%	Score	3894; DB 6; Length 784;
Best Local Similarity		92.7%	Pred. No.	0; Mismatches
Matches	706; Conservative	35;	Indels	0; Gaps
Qy	1	GPNCITRGVSSCQCLAVSPMCAMCSDAELPLGSPRCDLKENLKDCAPESTEFPSE	60	
Db	23	GPNCITRGVSSCQCLAVSPMCAMCSDAELPLGSPRCDLKENLKDCAPESTEFPSE	82	
Qy	61	ARVLEDPLSKGSQDSQVQSPORIALRLRDPDSKNISIQVROVEDYPVDIYMLD	120	
Db	83	ARILEAPLSPSKGSQDSQVQSPORIALRLRDPDSKNISIQVROVEDYPVDIYMLD	142	
Qy	181	KTTCIPLMFGYKHVLTLDQVTRFNEEVKKQSVSRNRDAPEGGDAIMQATVCDEKIGRN	240	
Db	203	KTTCIPLMFGYKHVLTLDQVTRFNEEVKKQSVSRNRDAPEGGDAIMQATVCDEKIGRN	262	
Qy	241	DASHLVLVFTDAKTHIALDGLLAGTVQPNQGCHGSDNHYASASTMYPQLGIMTEKL	300	
Db	263	DASHLVLVFTDAKTHIALDGLLAGTVQPNQGCHGSDNHYASASTMYPQLGIMTEKL	322	
Qy	301	QKNINLIFAVTENVLYQNSELIPGTVQPNQGCHGSDNHYASASTMYPQLGIMTEKL	360	
Db	323	QKNINLIFAVTENVLYQNSELIPGTVQPNQGCHGSDNHYASASTMYPQLGIMTEKL	382	
Qy	361	DLPEELSLSFNATCLNEVTPGKSMGKIGDVSISIEAKYRGCPQEKEKSTIPIVG	420	
Db	383	DLPEELSLSFNATCLNEVTPGKSMGKIGDVSISIEAKYRGCPQEKEKSTIPIVG	442	
Qy	421	FKDSLTVQVTFNCDCCQAOAEPDSPRCDNGTFECGVRCGPGWLGSQCESEEDYRP	480	
Db	443	FKDSLTVQVTFNCDCCQAOAEPDSPRCDNGTFECGVRCGPGWLGSQCESEEDYRP	502	
Qy	481	SQDDECSPREPGQPVCSORGCLCQCVCHSSDGFKITGYCECDDFSCVYKGCMSHG	540	
Db	503	SQDDECSPREPGQPVCSORGCLCQCVCHSSDGFKITGYCECDDFSCVYKGCMSHG	562	

QY	541 QCCGGDCLCDSDWITGTYCNCNTTRDTMSSNGLCSGRGKECCGCVCTQPGSQTCEK 600
Db	563 QCSCGDLCDSDWITGTYCNCNTTRDTMSSNGLCSGRGKECCGCVCTQPGSQTCEK 622
QY	601 CPTCPDACTFKKECKFDRGALIDENTNRRYCRDEIESVKELKDTGKDANVNTYKE 660
Db	623 CPTCPDACTFKKECKFDRGALIDENTNRRYCRDEIESVKELKDTGKDANVNTYKE 682
QY	661 DDCVWRQYEDSGKSILVYVEPCKPGFDLILVLLSYNGAIIIGLALLIWKLIT 720
Db	683 DDCVWRQYEDSGKSILVYVEPCKPGFDLILVLLSYNGAIIIGLALLIWKLIT 742
QY	721 IHDKEAKFEEERAKWDTANNPPLYKEATSTFTNTYRGT 762
Db	743 IHDKEAKFEEERAKWDTANNPPLYKEATSTFTNTYRGT 784
RESULT	6
Q9W16	PRELIMINARY; PRT; 723 AA.
AC	Q9W16; 01-MAY-2000 (TREMBlre. 13, Created)
DT	01-MAY-2000 (TREMBlre. 13, Last sequence update)
DT	01-DEC-2001 (TREMBlre. 19, Last annotation update)
DE	BETA 3 INTEGRIN, GPIIIA.
OS	Rattus sp.
OC	Eutheria; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
RN	{1}
RP	SEQUENCE FROM N.A.
RX	MEDLINE-93298675; PubMed=8512576;
RA	Cleutin A.M., Rosa J.P., Letourneur F., Ponczi M., Rifat S.;
RT	"A comparative analysis of cDNA-derived sequences for rat and mouse beta 3 integrins (GPIIIA) with their human counterpart."
RL	Biochem. Biophys. Res. Commun. 193:771-778(1993).
CC	-i- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (BY SIMILARITY).
CC	-i- PTM: THE CYSTEINE RESIDUES ARE INVOLVED IN INTRACHAIN DISULFIDE BONDS (BY SIMILARITY).
CC	-i- SIMILARITY: BELONGS TO THE INTEGRIN BETA CHAIN FAMILY.
DR	InterPro: IPR002086; Aldehyde_dehydr.
DR	InterPro: IPR00561; EGF-like.
DR	InterPro: IPR002369; Integrin_B.
DR	InterPro: IPR001169; Integrin_beta_C.
DR	InterPro: IPR002035; WWA.
DR	PFam: PF00362; integrin_B; 1.
DR	ProDom: PD001811; Integrin_B; 1.
DR	SMART: SM00001; EGF-like; 1.
DR	SMART: SM00187; INN: 1.
DR	SMART: SM00327; WWA; 1.
DR	PROSITE: PS00070; ALDEHDE_DEHYDR_CYS; UNKNOWN_1.
DR	PROSITE: PS00022; EGF_1; UNKNOWN_2.
DR	PROSITE: PS00186; EGF_2; UNKNOWN_1.
DR	PROSITE: PS00433; INTEGRIN_BETA; 3.
DR	Cell adhesion; Cytoskeleton; Extracellular matrix; Glycoprotein; KW
DR	Integrin; Repeat; Transmembrane.
SQ	SEQUENCE 723 AA; 80233 MW; B31D317E6ECF90F CRC64;
Query Match	86.8%; Score 3606; DB 11; Length 723;
Best Local Similarity	90.5%;
Matches	654; Conservative 42; Mismatches 27; Indels 0; Gaps 0;
QY	40 LKENLKDNGAPESTEFPSEARVLEDRDLSDKGDSQSYQVSPQIALRRLPDSK 99
Db	1 LKNUILKDNGAPESTEFPSEARVLEDRDLSDKGDSQSYQVSPQIALRRLPDSK 60
RESULT	7
Q92071	PRELIMINARY; PRT; 781 AA.
ID	Q92071; Q90829;
AC	Q92071; Q90829;
DT	01-NOV-1996 (TREMBlre. 01, Last sequence update)
DT	01-DEC-2001 (TREMBlre. 19, Last annotation update)
OS	INTEGRIN_BETA3
OC	Gallus gallus (Chicken).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi; Archosauia; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.
OC	Archosauia; Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.
RN	{1}
RP	SEQUENCE FROM N.A.
RC	STRAIN=WHITE LEGION;
RC	MEDLINE-94164000; PubMed=8119143;
RA	Mimura H., Cao X.P., Ross F.P., Chiba M., Teitelbaum S.L.;
RT	"1,25-dihydroxyvitamin D ₃ transcriptionally activates the beta 3-integrin subunit gene in avian osteoclast precursors."
RL	Endocrinology 134:1061-1066(1994).
RN	{2}
RP	SEQUENCE OF 1-26 FROM N.A.
RC	STRAIN=BONE MARROW;
RC	MEDLINE-940657; PubMed=8262978;
RA	Cao X., Ross P.F., Zhang L., Macdonald P.N., Chappel J., Teitelbaum S.L., Patrick F.;
RT	"Cloning of the promoter for the avian integrin beta 3 subunit gene and its regulation by 1,25-dihydroxyvitamin D ₃ ."

QY	319	QNSVELIPIGTIVGVLMSDSDNVLQLTIVDAYKIRSKVELEVRDPEELSNATCNNE	378
Db	241	QNSVELIPIGTIVGVLMSDSDNVLQLTIVDAYKIRSKVELEVRDPEELSNATCNNE	3000
QY	379	VIPGLKSCMGKIGDVSFESTEAKVRGCPQEREKSFIIKPGFEDSLIVQVIFDCACQ	438
Db	301	VIPGLKSCMGKIGDVSFESTEAKVRGCPQEREKSFIIKPGFEDSLIVQVIFDCACQ	3600
QY	439	AQAEPNPSHRCCNNGNTFEGCVCRCGPGCWLGSOCESCEFDYRPSQDECSPRECPQVPSQR	498
Db	361	AQAEPNPSHRCCNNGNTFEGCVCRCGPGCWLGSOCESCEFDYRPSQDECSPRECPQVPSQR	4200
QY	499	GECILGOCVCWCHSSDFGKITEGKYCECDDFSVCVYKGCMSHGQCSCDCCLCPSDWTGVCY	558
Db	421	GECILGOCVCWCHSSDFGKITEGKYCECDDFSVCVYKGCMSHGQCSCDCCLCPSDWTGVCY	4800
QY	559	NCTTRIDTCMSSNGLLCSGGRGKCECGGSCVCIQPGSYCDTCERKOPTCPDACTKEVCK	618
Db	481	NCTTRIDTCMSSNGLLCSGGRGKCECGGSCVCIQPGSYCDTCERKOPTCPDACTKEVCK	5400
QY	619	KFDRGALHDENTCNRYCRDDEIESVKEKLDTGKDAVNCTYKNEDDCVVRQYDSSGEI	678
Db	541	KFDRGALHDENTCNRYCRDDEIESVKEKLDTGKDAVNCTYKNEDDCVVRQYDSSGEI	6000
QY	679	LYVEEPECKPKDPDILVWLSYGAILLGIAALIWKLLITHDRKEFAKEERARK	736
Db	601	LYVEEPECKPKDPDILVWLSYGAILLGIAALIWKLLITHDRKEFAKEERARK	6600
QY	739	WDTANNPPLYKEATSTFTNIT	758
Db	661	WDTANNPPLYKEATSTFTNIT	680
RESULT	9		
RX	007012	PRELIMINARY;	PRT: 788 AA.
AC	007012;		
DT	01-NOV-1996 (TREMBLRE. 01, Created)		
DT	01-JAN-1999 (TREMBLRE. 09, Last sequence update)		
DT	01-DEC-2001 (TREMBLRE. 19, Last annotation update)		
DE	INTEGRIN BETA-3 SUBUNIT PRECURSOR.		
OS	Xenopus laevis (African clawed frog).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;		
OC	Xenopodinae; Xenopus; XENOBIL_TAXID=8355;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE:94040374; PubMed=7693527;		
RA	Ransom D.G., Hens M.D., Desimone D.W.;		
RT	"Integrin expression in early amphibian embryos: cDNA cloning and characterization of Xenopus beta 1, beta 2, beta 3, and beta 6 subunits".		
RT	Dev. Biol. 160: 265-275 (1993).		
RL	- - SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (BY SIMILARITY).		
CC	- - ITM: THE CYSTEINE RESIDUES ARE INVOLVED IN INTRACHAIN DISULFIDE BONDS (BY SIMILARITY).		
CC	- - SIMILARITY: BELONGS TO THE INTEGRIN BETA CHAIN FAMILY.		
DR	EMBL; LJ3591; RAA17427.1; -		
DR	InterPro: IPR02086; Aldehyde_dehydr.		
DR	InterPro: IPR00561; EGF-like.		
DR	InterPro: IPR02369; Integrin_B.		
DR	InterPro: IPR00169; Integrin_beta_G.		
DR	InterPro: IPR03659; PSI.		
DR	Pfam: PF00362; Integrin_B; 1.		
DR	PRINTS; PRO1186; INTEGRIN_B.		
DR	ProDom: PD00181; Integrin_B; 1.		
DR	SMART; SM00181; EGF; 1.		
DR	SMART; SM00187; INB; 1.		
DR	SMART; SM00423; PSI; 1.		
DR	PROSITE; PS00070; ALDEHYDE_DEHYDR_CYS; UNKNOWN_1.		

DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update).
 DE INTEGRIN BETA-5 SUBUNIT (FRAGMENT).
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
 Bovidae; Bovine; Bos.
 OC
 OX NCBI_TaxID=9913;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=MAMMARY GLAND;
 RA Andersen M.H.; Rasmussen J.T.; Berglund L.; Petersen T.E.;
 RY Submitted (OCT-2000) to the EMBL/GenBank/DDBJ databases.
 DR EMBL; AR317198; AAG38594.1; -
 DR InterPro; IPR00501; EGF-like.
 DR InterPro; IPR02369; Integrin_B.
 DR InterPro; IPR03059; PSI.
 DR InterPro; IPR02035; VWA.
 DR Pfam; PF00362; integrin_B; 1.
 DR PRINTS; PRO01186; INTEGRIN_B.
 DR PRODOM; PDO01811; Integrin_B; 1.
 DR SMART; SM00181; EGF; 1.
 DR SMART; SM00001; EGF like; 1.
 DR SMART; SM00187; INB; 1.
 DR SMART; SM00423; PSI; 1.
 DR SMART; SM00277; VWA; 1.
 DR PROSITE; PS00022; EGF_1; UNKNOWN_1.
 DR NON_TER 1
 FT SEQUENCE 791 AA; 87144 MW; 3FB45E62374169BC CRC64;

RESULT 11
 OGGLPO PRELIMINARY; PRT; 798 AA.
 ID OGGLPO
 AC OGGLPO
 DT 01-MAR-2001 (TREMBLrel. 16, Created)
 DT 01-MAY-2001 (TREMBLrel. 16, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE INTEGRIN BETA-1 SUBUNIT.
 GN CD29.
 OS Sus scrofa (Pig).
 OC Bovaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 OX NCBI_TaxID=9823;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20426040; PubMed=1097224;
 RA Jimenez-Marin A.; Garrido J.J.; de Andres-Cara D.F.; Morera L.,
 RA Barancho M.J.; Llanes D.;
 RT "Molecular cloning and characterization of the pig homologue to human
 CD29, the integrin beta1 subunit";
 RT Transplantation 70:649-655 (2000).
 DR EMBL; AF192528; AAG16757.1; -
 DR InterPro; IPR0030561; EGF-like.
 DR InterPro; IPR02369; Integrin_B.
 DR InterPro; IPR002035; VWA.
 DR Pfam; PF00362; integrin_B; 1.
 DR PRINTS; PRO01186; INTEGRIN_B.
 DR PRODOM; PDO01811; Integrin_B; 1.
 DR SMART; SM00187; INB; 1.
 DR SMART; SM00423; PSI; 1.
 DR SMART; SM00277; VWA; 1.
 DR PROSITE; PS00022; EGF_1; UNKNOWN_1.
 DR NON_TER 1
 FT SEQUENCE 798 AA; 88265 MW; 80773E985E38BA90 CRC64;

Query Match 55.7%; Score 2315.5; DB 6; Length 791;
 Best Local Similarity 54.8%; Pred. No. 8.4e-184; Mismatches 211; Indels 19; Gaps 9;
 Matches 426; Conservative 122; MisMatches 211; Indels 19; Gaps 9;
 QY 1 GPNICITTRGVSCQQCLAVSPMCAMCSDE--ALPLPGSPRDLKENLKKNCAPESIEFP 57
 1 |||||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:
 15 GLNICITSGSATSCCECLLHPIKRCAWCPKEDFGLRSVTSRCDLKENLKKNCAPESIEFP 57
 DR 58 VSEARVLDRPLDKSGGDS-SQTVTSPSPTIALRPLRPPDSKNSFIQVROVEDYVDIY 116
 DR 74 ASSTQYRLSLPLSKQSSPAGSPASDVIQLPQEVTVLRLPDRDTAQFOLQVROVEDYVDIY 133
 QY 117 IMLDSYMSKDDLWISLQNLGTQKATOMKLTSLNRLRGGAFVDFKPVSPYMISSPPEALEP 176
 DR 134 LMDLSLMSKDDLDNTRSLGKTLKAEEMRKLTSNFRFLGGFSRVDKNITSPSY-TAPRYQTNP 192
 QY 177 C--YDKKTTCTCPMEGKHKVHLTLDVOTRENEEVKQSVSRNRDAPEGGDAIMATVCD 234
 DR 193 CIGYKLFKPNCVPSGFRHLPLTDRVDSFNEEVKQVRVSRNRDAPEGGDAVLOAAYKE 252
 DR 235 KIGWRNDASHLVLVFTDAKTHALDGLAGIVQNDGOCHVGSIDNHYSASTMYPISGL 294
 DR 253 KIGWRKDHALHLVFTDAKTHALDGLAGIVQNDGOCHVGSIDNHYSASTMYPISGL 294
 QY 295 NTEKLSQKNJNLLIFAVTENVVNLQNYNSELPGTUVGVLSSMDSSNVLQDLYVAGKTRSK 354
 DR 313 LGEKLAAENNLLIFAVTKHNMLYMKNFTALPITVIEHLHDSKNIQLLITNAYNSIRSK 372
 QY 355 VELEYRDLPLPELSEFSNATCINNEVPIGKSKMQLKIGTVTSIEAKVRCGQEK EKS 413
 DR 373 VELSWDQEPDNLNLFPTATCQDGVSYPGORKCEGLKIGDTASLRVLSVBARSCPSKHYEHT 432
 DR 414 FTIKVGFKDSLWVQTFPCDCAQAOQAPNSHRNCNGTFCGVCRGGPMLGQSC 473
 DR 433 LALRIPVAFRDLSLEVGVTYNCRCGCSAGLEPDSSARCS-NGTVVCGLCBCNPYGLTRC 491
 QY 474 SEEDYRPSQODCSPREGOPVCSORGECUCQCVCHSSDFGKLTGKVCCECDPDSFCVRYKG 533
 DR 492 OEGESQSGYQNLNCREAEGPKLSCRGQSCNQSCFSEFEGKIGYSFCECDPNSCARNKG 551
 QY 534 EMC5GHHGQGSCGDCGLCDSDPWTGYYCNCCTRTDPMSSGNGLSSGRRGKCEGCGCVCTPGS 593
 DR 552 VLCSHGHECHGECKCHAGYIGDNCNCSTDISTCQARDGHICSDRGHCWVGQOCCTERGA 611
 QY 594 YGDTCEKOPTCDACTEKKVCEKKFDRGALDENTCNRYCRDE-IESVKELDTGKA 652
 DR 612 FGETCBKPCIPGPDACSTKRCVECLLHSSSSADQTCQNLCKDEVITRUDTIVKDDOEA 671
 QY 653 VNCTYKNEDDCVVRFOYVYEDSSGKSLVYVEPECPKGDILVVLSSNGAIIILLGAL 712
 DR 672 VLCFYKTAKDCVMMFTYSELPSGKSNLTVLREPECGTAPSAMTILLAVVGSILLTGFAEL 731
 DR 713 LIWKLIITIHKERKEAKFEEPERAKWDTANNPPLYKETSTFT-----NITYRG 762
 DR 732 VIKWLLVTHDRREFAKQFOSERSRARYEMASNPYLRKPISTHTVDFTFNKFENSYNGT 789

DT 01-MAY-1999 (TREMBLrel. 10, Created)
 DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE BIOTA INTEGRIN SUBUNIT
 OS *Biomphalaria glabrata* (Bloodfluke planorb).
 OC Eukaryota; Metazoa; Mollusca; Gastropoda; Pulmonata; Basommatophora;
 OC Planorbidae; Biomphalaria.
 OX NCBI_TaxID=6326;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99172885; PubMed=10072774;
 RA David B.J., Wu X.J., Yoshino T.P.;
 RT "Cloning of a beta integrin subunit cDNA from an embryonic cell line
 derived from the freshwater mollusc, *Biomphalaria glabrata*.";
 RL Gene 228:213-223(1999).
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (BY SIMILARITY).
 CC -1- PTM: THE CYSTEINE RESIDUES ARE INVOLVED IN INTRACHAIN DISULFIDE
 CC BONDS (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE INTEGRIN BETA CHAIN FAMILY.
 EMBL: AR060203; AAC67603; 1;
 DR InterPro: IPR00561; EGF-like.
 DR InterPro: IPR002369; Integrin_B.
 DR Integrin: Repeat; Transmembrane.
 DR PROSITE: PS00243; INTEGRIN_BETA_3.
 DR SMART: SM00187; INB; 1.
 DR PROSITE: PS00022; EGF_1; UNKNOWN_3.
 DR PROSITE: PS01186; EGF_2; UNKNOWN_1.
 DR SMART: SM00001; EGF-like; 1.
 DR PROSITE: PS00022; EGF_1; UNKNOWN_3.
 DR PROSITE: PS01186; EGF_2; UNKNOWN_1.
 DR SMART: SM00187; INB; 1.
 DR PROSITE: PS00022; EGF_1; UNKNOWN_3.
 DR PROSITE: PS00243; INTEGRIN_BETA_3.
 KW Cell adhesion; Cytoskeleton; Extracellular matrix; Glycoprotein;
 KW Sequence; Repeat; Transmembrane.
 SEQUENCE 788 AA; 87632 MW; 3F3JFAFE0848163 CRC64;

RESULT 14

Q9NAST7 ID Q9NAST7
 RP PRELIMINARY; PRT; 837 AA.
 AC Q9NAST7;
 RX DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 RA DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 RT DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE INTEGRIN BETA SUBUNIT PRECURSOR.
 GN BINT.
 OS Anopheles gambiae (African malaria mosquito).
 OC Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicidae;
 OC Anopheles; Anopheles gambiae (African malaria mosquito).
 OC Nematocera; Culicidae; Anopheles; Anopheles gambiae (African malaria mosquito).
 OX NCBI_TaxID=7115;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=G3;
 RX MEDLINE=21331701; PubMed=11437913;
 RA Machaiak V., Lycett G., Blasius C., Louis C.;
 RT "Beta-Integrin of Anopheles gambiae: mRNA cloning and analysis of
 structure and expression.";
 RL Insect Mol. Biol. 10:217-223 (2001).
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (BY SIMILARITY).
 CC -1- PTM: THE CYSTEINE RESIDUES ARE INVOLVED IN INTRACHAIN DISULFIDE
 CC BONDS (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE INTEGRIN BETA CHAIN FAMILY.
 EMBL: AR22975; CAC00620.1; -.
 DR InterPro: IPR00561; EGF-like.
 DR InterPro: IPR002369; Integrin_B.
 DR InterPro: IPR001169; Integrin_beta_C.
 DR InterPro: IPR003059; PST.
 DR InterPro: IPR002035; WWA.
 DR PROSITE: PS00362; Integrin_B; 1.
 DR PROSITE: PS01186; INTEGRIN_B.
 DR PRODOM: PD00181; Integrin_B; 1.
 DR SMART: SM00187; INB; 1.
 DR SMART: SM00423; PSI; 1.
 DR PROSITE: PS00022; EGF_1; UNKNOWN_2.
 DR PROSITE: PS01186; EGF_2; UNKNOWN_1.
 DR PROSITE: PS00243; INTEGRIN_BETA_3.
 KW Cell adhesion; Cytoskeleton; Extracellular matrix; Glycoprotein;
 KW Integrin; Repeat; Transmembrane.
 FT SIGNAL 1 25 POTENTIAL.
 FT CHAIN 27 837 AA; 92791 MW; 30F139D6924D78 CRC64;

Query Match 39.9%; Score 1658; DB 5; Length 788;
 Best local Similarity 42.7%; Pred. No. 4.3e-129;
 Matches 329; Conservative 136; Mismatches 267; Indels 38; Gaps 20;

QY 480 PSQQ-DPCSPREGQPYCSQRGECLGCCCVCH--SSDFGK-ITGKYCECDDFSCVRYKEM 535
 DB 503 SKDSIKCCTAPNATLPCSGRGSCVCGECCLCNTRSSDQAFTSGPYCECDYSCNQFQI 562
 QY 536 CSG -HQCSGDCLDDSDWGYCCTTRIDTCSSGLCSRGKCEGSCVQIOP 593
 DB 563 CGPERKRCNCGKCECKLGNYTACEEELSMEMKCTTDKGGLVCGHNGCUGKGKCVG--ENQ 620
 QY 594 Y-GDTEECKPCPQDACTFKKECCKFDRGALHDENTCNRYCDEI--ESVKELKDG 649
 DB 621 YTGFKEQQCPACDPCDQKLELYPCWQCKAKFGKQKLIQER-CIRECMEIEFHDKIRE--G 675
 QY 650 KDAVNTYKNEEDDCVYVRFQYEDSSGKSLIYVVEPECPCPKGPDILVVLSSVMAILLGL 709
 DB 676 HGIQCPFKDQDCWAVYTYEYDRGKVKQAITKVCVFQDQLNLAIVGUVGIVAVGL 735
 QY 710 AALIWKLLITHDRKEFAFEERARKADTANNPLKETSTFTNTY 759
 DB 736 FLILIWKLLITHDRKEFAFEERARKADTANNPLKETSTFTNTY 785

Query Match 38.7%; Score 1609; DB 5; Length 837;
 Best local Similarity 39.5%; Pred. No. 5.6e-125;
 Matches 323; Conservative 138; Mismatches 284; Indels 72; Gaps 19;

QY 4 ICTTRGVSSCQCLAVSPMCWCSDAELP-LGSPRCDEKENLKDNCAPESTEPVSEAR 62
 RA :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||

RA Marsden M., Burke R.D.,
 RL Submitted (JAN-197) to the EMBL/GenBank/DBJ databases.
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (BY SIMILARITY).
 BONDS (BY SIMILARITY):
 -1- SIMILARITY: BELONGS TO THE INTEGRIN BETA CHAIN FAMILY.
 DR EMBL; U77584; ARB39799; 1.
 DR EMBL; 077587; ARB39741; 1.
 DR InterPro; IPR00561; EGF-like.
 DR InterPro; IPR02369; Integrin_B.
 DR InterPro; IPR01169; Integrin_beta_C.
 DR InterPro; IPR02035; wWA.
 DR Pfam; PF00322; Integrin_B; 1.
 DR ProDom; PD001811; Integrin_B; 1.
 DR SMART; SM0001; EGF-like; 1.
 DR SMART; SM00187; INB; 1.
 DR SMART; SM0023; PSI; 1.
 DR PROSITE; PS00022; EGF_1; UNKNOWN_2.
 DR PROSITE; PS01186; EGF_2; UNKNOWN_1.
 DR PROSITE; PS0243; INTEGRIN_BETA; 2.
 KW Cell adhesion; Cytoskeleton; Extracellular matrix; Glycoprotein;
 KW SEQUENCE: 783 AA; 85530 MW; BB045C6FD88FB2 CRC64;

Query Match Best Local Similarity 38.2%; Score 1588.5; DB 5; Length 783; Matches 317; Conservative 144; Mismatches 51; Gaps 21;

QY 12 SCQOCLAVSPMCWC---SDEALPLGSPRCDEKENLKDNC---PESIEPVSEAR 62
 DB 35 TCGECISLSDSCGWCNTLNNTD---GNPOCDLASSLQRGCSQIVDPDS-----TM 84
 QY 63 VLEDRPLSDKGSDQSQQVYQSPQALRRLRPPDRDKNFSQVQROVEDYVYDYLMDSY 122
 DB 85 VLGWTSNLSAGAQPQGQVQYQVPRQYDPLKRGKPVWMLPGPAAEDYVYDYLMDSY 144
 QY 123 SMKDDIWSQNLIGTLKATQRKLTSLNRGFLGGAFVDPKQVSPYMTSPPELENPCYDMK 182
 DB 145 SMEDDLISKMLDGLTASERKNTNSFRFLGGFFGSDVKTWVYDPIV-SIVPEBKLIAC---T 199
 QY 183 TCLPMFPGYKVLTDQVTHNEEVKKQSVSRNDAPEGGFDALMOATWDEKIGWRDA 242
 DB 200 GCEAPYGFKNVPLPLNENTLFSETVMNQRAASNQGLDAPEGMDALMQITVQGNOJGWRENA 259
 QY 243 SHLVIFFTDKTHALGRLAGIYQVNDQCHVGSS-DNHYASITMDPSLGLMTEKLSQ 301
 DB 260 RHLVYPTDSS---IMPETKGGITPNDQCYLPSMSDNVQLLVDAYGKIRSKVELERD 361
 QY 302 KNINLIFAVNVNLYQNSLEIPTGTVGLSMSDNVQLLVDAYGKIRSKVELERD 377
 DB 319 NSVPIFAVQKEFELYNNIYQIEGATAGILAOQDNNTVQYQDNKNTSVEV 377
 QY 362 LPEELSLSFNATCLNNENFVTPGLKSCMGKLTGDFVSFSSTEAKVRCGCPDEKEKSFPIKPGF 421
 DB 378 APENVTIDGPHCPGQVTFGSGQVCEGLQQLGQDFTVNTFTITATGCPNNQYQQTVRPL- 436

QY 422 KDSLIVQVFFDCDCAQOQAPNHRCCNGCTFECCVRCGCGWLGSGQCESEEDYRS 481
 DB 437 QRELKVNFACDCDCEAQKVNQSCGNGTLECGSCICNPGHYRCGECSSDD-PT 494
 QY 482 QDDE---CSFREGQVPCSORGECLCGQVC---HSSDFKRTGNYCECODEFVYKGC 536
 DB 495 LEDNDAPCDTHTKHTRCGSGRCGCGCICRPNPSEVSGVSCFCECDNFENDRYLGLC 554
 QY 537 SG---HGQCSGGD-----CLCDSDWGTGYCNCITRTDQMSNSNGLCSCGRCECCSCV 588
 DB 555 GGSQRGQVQDETRRSQCRRCRGTYGDAECSTRVDTMIGL-TICNGEGVYCICGCK 613
 QY 589 IQPGSY-GDTECKPCTCDATCPTKKEFVCEKFKDFRGLHDENWCNRVCRDPEIRESWELK- 646

[1]
 RP SEQUENCE FROM N_A.

Db 614 NAGSSYRGALCQDPCTCSQCSRNEECYQCKAFGTGL--SKAOCDK-CPPPVIMVDNLEI 670
QY 647 DTGKIAVNCTYKNEEDCVRPQYEDSSGSKSLVVEPECPKGPDILVVLISVMGATL 706
Db 671 PTGSE--RCIADEDCCSIFTYAKSANALILYQKERVCFEPVDIMHVICIVVGII 728
QY 707 IGLAALLIWKLLITIHDKEFAFEERRAKWDTANPPLYKEATSTFTNTYR 760
Db 729 VGLALLLVWRLVYQDSREFASFEKERAGTHQONENPIYKPSISTFKNPTQ 782

Search completed: May 19, 2002, 12:26:43
Job time: 6465 sec